STOCHASTIC METHODS IN MODELING THE IMMUNE RESPONSE

by

Saishuai Tang

B.S. in Information and Computational Science, Fudan University, 2005

M.A. in Mathematics, University of Pittsburgh, 2007

Submitted to the Graduate Faculty of

the Department of Mathematics in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH

SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Saishuai Tang

It was defended on

July 18, 2010

and approved by

Professor David Swigon, Department of Mathematics

Professor Gilles Clermont, Critical Care Medicine

Professor Brent Doiron, Department of Mathematics

Professor Jonathan Rubin, Department of Mathematics

Dissertation Director: Professor David Swigon, Department of Mathematics

STOCHASTIC METHODS IN MODELING THE IMMUNE RESPONSE

Saishuai Tang, PhD

University of Pittsburgh, 2010

We discuss the application of deterministic and stochastic modeling techniques to problems in immunology. First, we employ the example of a host response to influenza virus infection to illustrate differences in the dynamical behavior of the deterministic and the stochastic models, and employ both versions in the analysis of the role of immune response in controlling and suppressing the infection. Second, we develop a dynamical model of vocal fold inflammation and use random sampling techniques to calibrate the model against available data.

In the stochastic model, we analyze three solution techniques - Gillespie's stochastic simulation algorithm, numerical solution of mean extinction time through the Laplace transform of the master equation, and approximate solution of a limiting case. Gillespie algorithm is capable of dealing with large systems as the required memory depends linearly on number of species, but is limited by computational time. Laplace's method is efficient and accurate but is limited by system size. We construct a novel combination of the two that takes advantage of both. We also derive an approximate Markov chain of the system and analytically compute the extinction times and probabilities in the limiting case when an infected cell generates large number of viruses. In the analysis of the human immune response, we find that innate response substantially reduces cell extinction probability, cellular response increases the virus extinction probability in limiting case, and adaptive response, combined with the other two, almost eliminates cell extinction and significantly increases virus extinction probability.

In the model of vocal fold inflammation, a four-variable ordinary differential equation

(ODE) model is presented. The ODE model characterizes cytokine interactions in phonotrauma and is calibrated with empirical data. Parameter values are estimated and their probability densities are sampled using Metropolis and parallel tempering algorithms. Sensitivity analysis showed that 6 of 17 parameters suffice to retain the sensitivity of the model trajectories. The reduced parameter set is applied to individual data to be calibrated and predict the individual outcome. The model is a part of larger study intended to find optimal treatment strategy for phonotrauma through a personalized vocal exercise or rest program.

Keywords: Stochastic Simulation Algorithm, Gillespie's Direct Method, Markov Chain Analysis, Numerical and Analytical Solution of Master Equations, Extinction Probability, Mean Extinction Time, Roles of Immune Components, Influenza Virus Infection, Vocal Fold Inflammation, Parameter Estimation and Reduction, Sensitivity Analysis, Metropolis Sampling, Parallel Tempering, Personalized Treatment Development.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I owe my gratitude to all the people who have made this dissertation possible and because of whom my graduate experience has been one that I will cherish forever.

First and foremost, I would like to thank my advisor, Professor David Swigon, for his constant support, guidance and patience to this fledgling researcher. Professor Swigon is kind, intelligent, and is always available for help and advice. Without his extraordinary insight and expertise, this dissertation would have been a distant dream.

I would also like to thank my research mentor, Professor Gilles Clermont, for his support, guidance and constructive discussion along the way.

I would also like to thank my other two committee members, Professor Jonathan Rubin and Professor Brent Doiron, for their suggestions and invaluable time to review this dissertation.

It is also my pleasure to thank Dr. Nicole Yee-Key Li and Dr. Qi Mi for their contribution and comments to my models. Especially thank Nicole, a superb biologist, for her efficient communication, patient support to this junior mathematician.

Then, I want to express my thanks to Mr. Ranqi Zhu, my boyfriend, for his constant encouragement and accompaniment all these years.

Last but not least, I am deeply grateful to my parents for their love, support, and allowing me, their only child, to pursue her education thousands of miles away from home.

1.0 INTRODUCTION

In the first part of this thesis, we discuss a model of the immune response to influenza infection. The human immune response to influenza A virus infection follows three distinct strategies aimed at reducing the level of the virus and the amount of damage to the system through negative feedback. The first strategy is the innate immune response, which renders a portion of epithelial healthy cells resistant to infection. The second strategy is the cellular immune response, which produces effector cells that destructs virus infected cells. The third strategy is the adaptive immune response, which produces antibodies that inactivate the virus particles. Each of the strategies represents a negative feedback on the virus production.

Mathematical modeling has proven to be a valuable tool in understanding of immune response to influenza infection, primarily with ODE models [3, 40]. In the deterministic model, the dynamics of the system are entirely predictable if sufficient knowledge of the system states is given. But due to the stochastic nature of the system, this prediction might not be true in certain situations. A stochastic model is a tool for estimating probabilistic potential outcomes by allowing for random variation in the inputs over time. It builds randomness into the simulation and therefore provides a better representation of real life. This thesis is concerned with the development of a stochastic model of the immune response. We first investigate a basic system of viral reproduction in the absence of any immune component. We design a stochastic model of the system and compare its behavior with a traditional ODE model. Three methods are utilized in analysis of the stochastic model: Gillespie's stochastic simulation algorithm [30], which is an exact procedure for simulating the time evolution of a well-stirred chemically reacting system, solution of mean extinction time through the Laplace transform of its master equation [67], which provides us explicit formulas of mean extinction time, and approximate solution in the limiting case of large

infectivity, which presents the analytical solutions when a single infected cell generates a large number of viruses. We then extend the stochastic model to include various immune system components. Our goal is to investigate the role of each feedback by studying its effect on the global stochastic dynamics of the virus-cell system.

In the second part of this thesis, we discuss a model of vocal fold inflammation. When people speak or sing, vocal folds are exposed to nearly continuous bio-mechanical stress - the vocal folds oscillate at a relatively high frequency and typically make contact at some point during the vibratory cycle. When the phonation conditions are phonotraumatic, increased intra-fold contact stress associated with certain voicing patterns can result in structural damage to the vocal fold mucosa [80, 37, 38]. If this condition persists without adequate rest, the persistent stress can further lead to tissue disorganization [36]. Traditionally, behavioral voice treatment involves complete or partial voice rest with the aim that the ensuing inflammatory response will subside on its own. However, recent data derived from concentrations of inflammation mediators in laryngeal secretions suggests that some forms of vocal fold tissue mobilization, i.e. resonant voice, may be able to modulate the inflammatory and healing process [53, 51].

In this thesis, a dynamic model is developed to characterize the process of vocal fold inflammation after phonation loading. Ensemble modeling technique is employed to estimate the parameter values and sample their probability densities. The number of independent parameters of the model is reduced through sensitivity analysis. We also conduct parameter reduction and employ the reduced model in prediction of the individual treatment outcome. The research reported here is a part of a larger project aimed in developing tools to help clinicians to decide which treatment benefits individual patients with acute or chronic phonotrauma through a personalized vocal exercise or rest program.

2.0 IMMUNE RESPONSE TO INFLUENZA INFECTION

2.1 INTRODUCTION

2.1.1 Deterministic and Stochastic Models

Systems biology employs primarily two types of models: deterministic and stochastic. In a deterministic model, the dependent variables usually represent concentrations of cells or chemicals and a set of ODEs or PDEs governs the evolution of the system. Such models tend to be accurate in the limit of large system size for well-mixed conditions. In a stochastic model, variables represent numbers of molecules or cells, and the evolution of the system is represented by a stochastic process described by a master equation. One important difference between the two models is the presence or absence of absorbing states. There are situations in which the deterministic model predicts that a population approaches a positive stationary level, while the stochastic model predicts that an extinction will occur with certainty. In such situations, one may ask the following questions:

- How to determine the probability of extinction and characterize the time required for the extinction?
- Does a bifurcation in the deterministic model correlate with a change in the behavior of the stochastic model?

The mean extinction time can be found if we can solve the master equations, but this is known to be surprisingly difficult. The analysis of the mean extinction time in the context of population class models has been discussed by several authors. Newman [67] and Nasell [66] presented the analysis of mean extinction time of a stochastic version (logistic stochastic model) of the Verhulst deterministic model [87] - a finite-state birth and death process

with an initially modest number of individuals, no immigration nor emigration, and one absorbing state. Vellela and Qian [83] presented a model with 3 species and 3 reactions with one absorbing state, but they enumerated the states according to one species, so it is still equivalent to the one dimensional logistic stochastic model. Nasell [65] also presented the mean extinction time analysis in a two-dimension model - the stochastic version of the SIR model (a model with three classes, susceptible, infected, and recovered, in a fixed population) and found that the mean extinction time is an exponential distribution with a finite expected value. There are also other discussions of the mean extinction time in the stochastic SIR and SIS (a model derived from the SIR model by simply considering that the individuals recover with no immunity to the disease, that is, individuals are immediately susceptible once they have recovered) models by Nasell [64], Dykman [24] and Kamenev [46]. In the two-dimension case, most of these papers analyzed the extinction time approximately in different parameter regions through a variety of mathematical techniques. Nasell [66] states that it is impossible to find explicit expression for the time to extinction of a stochastic model whose deterministic counterpart is nonlinear. Few papers discuss the extinction probability. Herwaarden [82] derive an asymptotic expression for the disease extinction probability at the end of a major outbreak in the two-dimension stochastic SIR model. The expression is derived by asymptotically solving a boundary value problem on the Fokker-Planck equation of the system.

In this thesis, we present a stochastic model derived as a model of virus dynamics in a host individual. First, we simulate system behavior through Gillespie's direct method, then we derive formulas of extinction time moments through the analysis of the Laplace transform of the master equation [67]. The Laplace transform converts the nonlinear master equation into a system of linear algebraic equations, which we can solve numerically to get the mean and variance of the extinction time. The model can also be approximated by a reduced Markov chain in one dimension with two absorbing states. We derive analytical expressions of both mean extinction time and probability of such a reduced chain.

2.1.2 In-host Virus Dynamics

In mathematical modeling of virus dynamics, ODE models of within-host viral infections of target cell populations were developed by Perelson and Nelson [71], Nowak and May [69] and others [61, 68, 72]. These models were primarily derived for HIV infection, and later applied to other viral infections [69]. They focused on the disease dynamics within an infected individual and provided the quantitative understanding of the level of virus production during infection [50, 72].

Leenheer and Smith [50] presented a global analysis of virus infection in an ODE model. In the absence of any immune component, the system follows a limit cycle trajectory that traverses through a region of almost complete absence of healthy cells, marked by a high proportion of infected and damaged cells. Bifurcation analysis of a complete immune response model, developed by Hancioglu [40], shows that the innate response, which renders a portion of healthy cells resistant to infection, eliminates this limit cycle and, instead, causes the system to converge to a chronic state with finite proportion of healthy and infected cells and virus level. The cellular response, which produces effector cells to remove the virus infected cells, lowers the magnitude of damage in the chronic state and, for some virus concentrations, causes the trajectory to pass very close to the healthy state. Finally, the adaptive response, which produces antibodies to neutralize viruses, stabilizes the healthy steady state as a global attractor of the system. (Unpublished results of M. Goldrich and D. Swigon).

Here we analyze the role of each immune component using a stochastic model, with the following questions in mind:

- Is any type of the immune component, by itself, able to clear virus and stabilize the healthy state?
- Can any type of the immune component be eliminated or weakened without disrupting the function of the immune system?
- Can any type of the immune component be enhanced and take over another type's role?

If the answer is negative to all these questions, we must conclude that the immune system is optimally designed for clearance of viral infection. To assess the effect of individual components of the immune system, we shall focus on two variables - the total damage to

the system, expressed as the proportion of cells incapable to perform their function, and the total amount of virus in the system. We shall also be interested in finding out how the systemic response depends on parameters describing the strength of the virus and the strength of each immune strategy.

2.2 BASIC SYSTEM

In the absence of immune response, the system consists of three components - the free virus particles (**V**), the healthy (susceptible) cells (**H**) and the infected cells (**I**). The most appropriate model of viral infection is a combination of a predator-prey model (with healthy cells as a prey and virus as the predator) and SIS epidemiological model. Neither of these models is correct in isolation. The classical predator-prey model predicts an unlimited growth of the prey in the absence of a predator, although in reality the amount of healthy cells should not exceed a limiting threshold. The SIS model would describe the situation in which the rate of infection of new individuals (cells) is proportional to the amount of infected cells, while during viral attack it should be proportional to the amount of free virus.

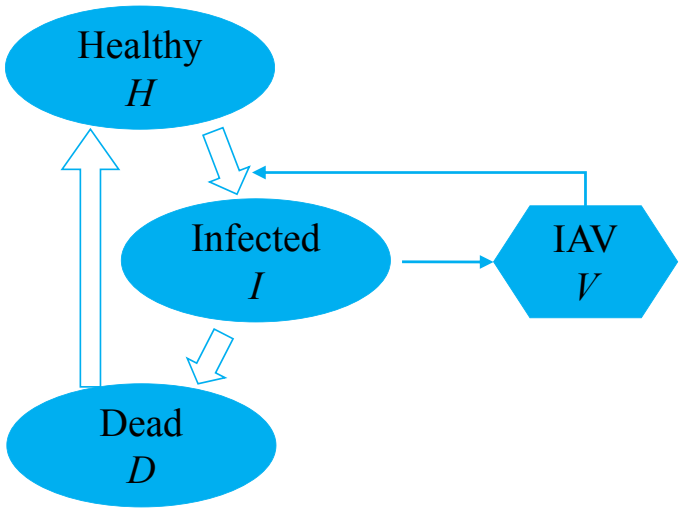

Figure 1: Basic System

2.2.1 ODE Model

ODE model based on the diagram in Figure 1 can be presented by the following system of differential equations:

$$\dot{V} = \gamma_V I - \gamma_{VH} HV - a_V V \tag{2.2.1}$$

$$\dot{H} = b_{HD}(1 - I - H)H - \gamma_{HV} VH \tag{2.2.2}$$

$$\dot{I} = \gamma_{HV} HV - a_I I \tag{2.2.3}$$

Here, the virus particles are produced in infected cells with about $n = \gamma_{VH}/a_I$ particles released upon the death of a single infected cell. The virus particles infect healthy cells and turn them into infected cell, with circa $m = \gamma_{VH}/\gamma_{HV}$ virus particles necessary to infect a single health cell. The virus level also naturally decreases with $k = a_V/\gamma_{VH}$ being the average amount of viral particles decayed before one infected cell is produced. All the variables are re-scaled by the homeostatic level $H^* = 10^{12}$ cells. H and I are given as proportions of H^* and hence range from 0 to 1, and V is given as the amount of viral particles per infected cell. The healthy cells regenerate with a rate proportional to the quantity $D = 1 - I - H$, which may be called the damage.

Equation (2.2.2) describes that the regeneration of healthy cells depends on the system being modeled. In the case of influenza, we use the following reasoning: The cells being infected by influenza are epithelial cells of the respiratory tract. The human body contains regulatory mechanism that strives to maintain a homeostatic level of such cells. Both the infected and healthy cells count toward this level and hence healthy cells will regenerate only if the total amount is below this homeostatic level. Other possibilities include a simpler logistic term $b_{HD}(1 - H)H$ or a more sophisticated $b_{HD}(1 - D - I - H)H$ in which dead cells are removed at a constant rate.

The analysis of influenza in humans [40] yielded the following estimates for the parameters of the dimensionless model: $\gamma_V = 510$, $\gamma_{VH} = 1.02$, $a_V = 1.7$, $b_{HD} = 4$, $\gamma_{HV} = 0.34$, $a_I = 1.5$. (In the stochastic model described later, we shall assume for simplicity $\gamma_{VH} = \gamma_{HV} = 0.7$ and hence $m = 1$.) These constants correspond to roughly 340 ($n = \gamma_V/a_I = 510/1.5 = 340$)

virus particles released from an infected cell per day and one infected cell produced after about 5 ($k = a_V/\gamma_{HV} = 1.7/0.34 = 5$) virus particles decay.

The parameter γ_V describes the reproductive ability of the virus and is the main virus-dependent parameter affecting the behavior of the system equation (2.2.1). Figure 2 and 3 show the dependence of the behavior of the system on γ_V.

When γ_V is less than the critical value $\gamma_V^{BP} = (\gamma_{VH} + a_V)a_I/\gamma_{HV}$, then the system has a stable equilibrium state $(V, H, I) = (0, 1, 0)$, which we call *healthy state*, and an unstable equilibrium state $(V, H, I) = (0, 0, 0)$, which we call *dead state*. When γ_V is above γ_V^{BP}, the healthy state becomes unstable and the system has an additional steady state, which we call *chronic state*. The chronic state is stable and is characterized by nonzero values of V and I that depend on γ_V:

$$H_C = \frac{a_V a_I}{\gamma_V \gamma_{HV} - \gamma_{VH} a_I} \tag{2.2.4}$$

$$V_C = \frac{b_{HD}(1 - H_C)}{\gamma_{HV}(1 + b_{HD} H_C/a_I)} \tag{2.2.5}$$

$$I_C = \frac{\gamma_{HV}}{a_I} H_C V_C \tag{2.2.6}$$

In equation (2.2.4), $\gamma_V \gamma_{HV} - \gamma_{VH} a_I > \gamma_V^{BP} \gamma_{HV} - \gamma_{VH} a_I = a_V a_I$. Thus we have $0 < H_C < 1$. When γ_V above γ_V^{HB}, the chronic state becomes unstable and a limit cycle appears through a Hopf bifurcation. The limit cycle persists in the limit as $\gamma_V \to \infty$, but approaches exponentially close to the dead state, which remains unstable for all values γ_V. See the bifurcation diagram in Figure 2 and Figure 3.

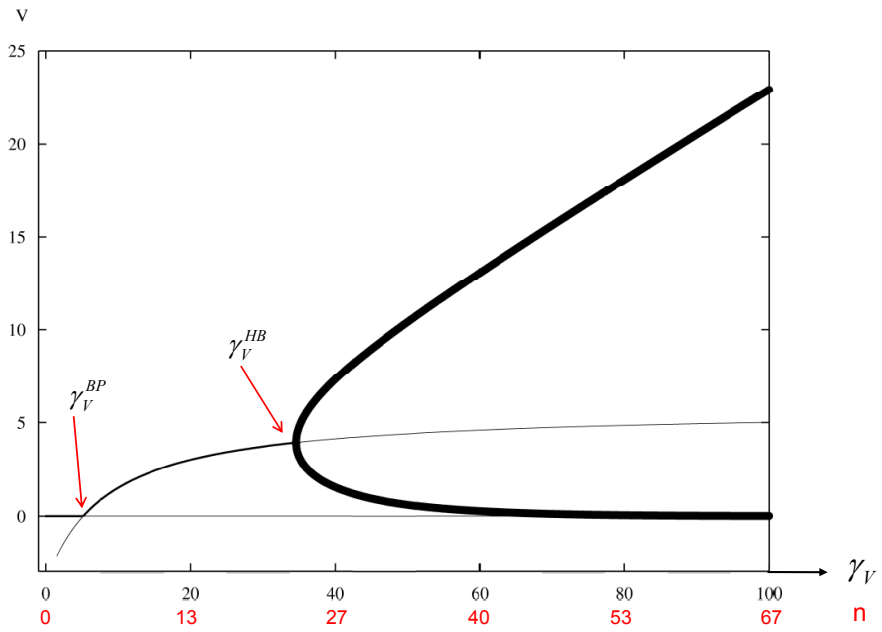

Figure 2: Bifurcation diagram of the basic system - virus dependence on γ_V and n. $n = \gamma_V/a_I$ represents the number of viruses generated by a single infected cell. The system has different behaviors through 3 parameter regions.

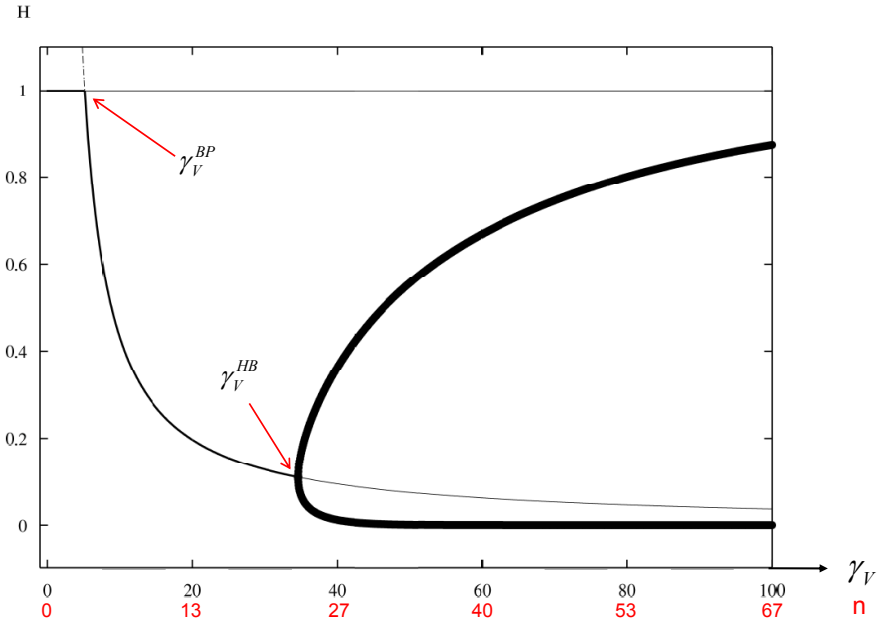

Figure 3: Bifurcation diagram of the basic system - healthy cell dependence on γ_V and n.

A standard course of infection starts in the vicinity of the healthy state with no infected or dead cells and a small amount of the virus. We typically use $(V(0), H(0), I(0)) = (0.03, 1, 0)$ as the initial state in the ODE model. In the limit cycle regime, this trajectory rapidly approaches a limit cycle that can be divided roughly into three periodically repeating parts: (i) rapid growth of the number of infected cells and the number of viral particles until exhaustion of healthy cells, followed by (ii) decay of infected cells with slowing production and decay of the virus, (iii) regeneration of the healthy cells.

One important unrealistic feature of the basic ODE model is that the system is able to make a recovery even after almost complete absence of healthy cells. This result is an oversimplification: if we considered the cells to be a part of an organism, then the absence of healthy cells would result in a systemic failure and death of the organism. In our case we may pronounce the system dead if the number of healthy cells drops below a certain threshold, say 0.1. Of course, the choice of that threshold will affect the outcome. This problem results from the stochastic nature of the underlying problem and can be addressed in the context, which we do below.

2.2.2 Stochastic Model

The set of interactions between free virus particles (**V**), the healthy (susceptible) cells (**H**) and the infected cells (**I**) can also be written conveniently as a system of chemical reactions that symbolically describes the interactions between the components. The reaction network of the basic system is:

$$V + H \xrightarrow{\gamma_{VH}} I \tag{2.2.7}$$

$$I \xrightarrow{a_I} nV + D \tag{2.2.8}$$

$$V \xrightarrow{a_V} \varnothing \tag{2.2.9}$$

$$H + D \xrightarrow{b_{HD}} 2H \tag{2.2.10}$$

A single infected cell produces n numbers of viruses. $n = \gamma_V / a_I$ can be derived from (2.2.1) and (2.2.3). The healthy cells regenerate with a rate proportional to the quantity $D = H^* - I - H$. These chemical reactions give rise to the system of equations (2.2.1) $\sim$

(2.2.3) under the assumption of mass-action kinetics, with one exception: for simplicity we assume that the rate γ_{HV} is equal to γ_{VH} and hence $m = 1$. (Distinct values can be achieved by adding the reaction $V + H \xrightarrow{\gamma_{HV} - \gamma_{VH}} H$ describing inefficient infection.)

The reactions (2.2.7) $\sim$ (2.2.10) imply that the total number $H^* = H + D + I$ of cells is invariant and hence the phase space is a semi-infinite triangular prism bounded by the planes $H = 0$, $I = 0$, $H + I = H^*$, and $V = 0$, with $D = H^* - H - I$. Both the healthy state $(V, H, I) = (0, H^*, 0)$ and the dead state $(V, H, I) = (0, 0, 0)$ are absorbing states. The dead state $(0, 0, 0)$ lies in the invariant facet of the phase space for which $H = 0$ - the green plane in Figure 4. Once a trajectory enters this facet, it will converge toward $(0, 0, 0)$, and this convergence will be monotonic in I . The healthy state $(0, H^*, 0)$ lies on the invariant line $V = I = 0$, the red line in Figure 4 - once a trajectory enters this line, it will converge toward $(0, H^*, 0)$, and this convergence will be monotonic in H.

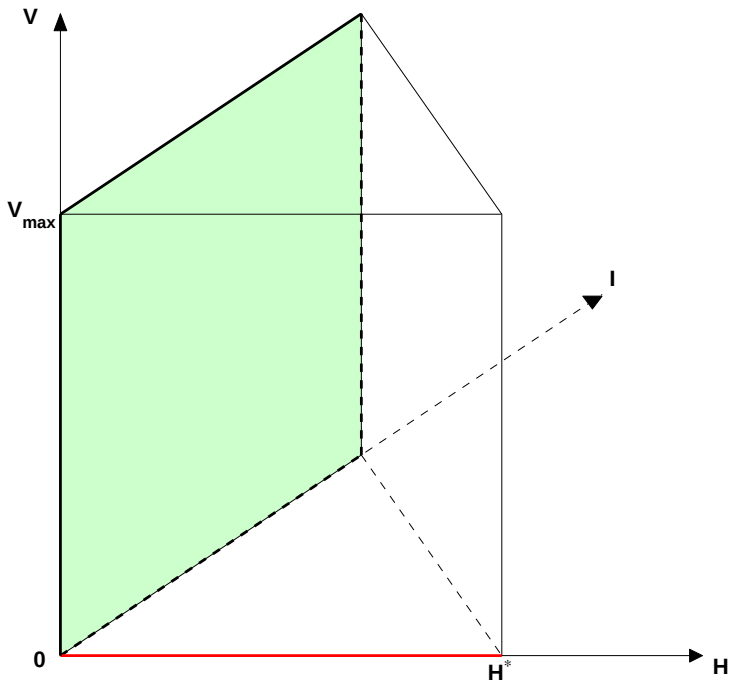

Figure 4: V-H-I triangular prism. The red line is the absorbing region for healthy state and green plane is the absorbing region for dead state .

The master equation for the reactions (2.2.7) $\sim$ (2.2.10) is given by:

$$\begin{aligned}\frac{d}{dt}P_{V,H,I}(t) &= \gamma_{VH}(H+1)(V+1)P_{V+1,H+1,I-1}(t) + a_I(I+1)P_{V-n,H,I+1}(t) \\ &+a_V(V+1)P_{V+1,H,I}(t) + b_{HD}(H-1)(H^*-H-I+1)P_{V,H-1,I}(t) \\ &-(\gamma_{VH}HV + a_II + a_VV + b_{HD}H(H^*-H-I))P_{V,H,I}(t)\end{aligned} \tag{2.2.11}$$

This equation is actually a system of ordinary differential equations that describes the occupancy probability for different states of a chemical system during a Markov process. It is also called a Kolmogorov equation. Each molecular species in the chemical model adds a dimension to the state space. Unfortunately, the master equation (2.2.11) is too complex to solve explicitly. So we have solved (a) the mean and limiting extinction probability of the system; (b) the conditional mean, variance and coefficient of variation of extinction times, of both viruses and healthy cells through the following method:

- Exact stochastic simulation - simulate the reaction network through Gillespie's direct method [30, 31].
- Numerical solution of the moments of extinction times through the Laplace transform of the master equations [67].
- Approximately analytically solution of the master equation in the limit of $n \to \infty$.

2.2.2.1 Stochastic Simulation Algorithm The most common method for solving a master equation in systems biology is a stochastic simulation algorithm (**SSA**) [30, 31, 29]. SSA is an essentially exact procedure for numerically simulating the time evolution of a well-stirred chemically reacting system. It is derived rigorously based on the same microphysical premise that underlies the chemical master equation [60, 32]. So it is logically equivalent to the master equation. The simulation statistically realizes the underlying Markov process. The trajectories of species in the model are simulated by randomly applying reactions. Many statistical properties of the system, such as the extinction time moments, probability distribution, etc., can be computed through SSA.

Gillespie [30] proposed two mathematically equivalent methods for implementing the SSA, namely the direct method (**DM**) and the first reaction method (**FRM**). DM is much

more efficient than FRM and it the most commonly employed method. A third formulation of SSA is the next reaction method (**NRM**) by Gibson and Bruck [29] published in 2000. NRM is an extension of FRM and it is significantly faster than FRM. It is also faster than DM under certain conditions - when many species and reaction channels are involved and the system is loosely coupled. But this statement is argued by Cao and his co-workers [10] who claimed that NRM was more efficient than the DM only in a very specialized class of problems. Simulation methods that are more efficient, but not exact, have been recently proposed, such as explicit [35, 8, 9] and implicit [75] τ-leaping method, optimized direct method [10]. A recent paper of Gillespie [34] contains a review of the relevant work in this fields in the past 30 years. Here, we illustrate the DM algorithm and briefly discuss the other methods.

Suppose we have M reactions, R_1, R_2, ..., R_M, in total. Define the following variables at time t [30, 47].

- C_μ: the rate of reaction R_μ.
- h_μ: the number of reactant combinations for reaction R_μ.
- $a_\mu dt = h_\mu C_\mu dt$: probability that reaction R_μ occurs in the next time interval dt.
- $a_0 dt = \sum_{\mu=1}^{M} a_\mu dt$: the sum of the average probabilities that any reaction occurs in the next time interval dt
- $P_0(\tau)$: probability that no reaction occurs during $(t, t+\tau)$
- $P(\tau, \mu)d\tau = P_0(\tau)h_\mu C_\mu d\tau$: probability at time t that the next reaction is R_μ and occurs in the interval $(t+\tau, t+\tau+d\tau)$. So $P(\tau, \mu)$ is the reaction probability density function.

For example, if reaction R_1: $A + B \to AB$ occurs and the number of molecules A and B are X and Y, then $h_1 = XY$ and $a_1 = XYC_1$.

Divide the time interval $(t, t+\tau)$ into K subintervals with width ε. The probability that none of the reactions occurs in the N subintervals is:

$$P_0(\tau) = \big(1 - \sum_{i=1}^{M} h_i C_i \varepsilon\big)^K = \lim_{K\to\infty} \Big(1 - \frac{\sum_{i=1}^{M} h_i C_i \tau}{K}\Big)^K = e^{-\sum_{i=1}^{M} h_i C_i \tau} = e^{-a_0 \tau} \qquad (2.2.12)$$

Then

$$P(\tau, \mu) = P_0(\tau) h_\mu C_\mu = e^{-a_0\tau} a_\mu = (a_0 e^{-a_0\tau}) \left(\frac{a_\tau}{a_0}\right) = P(\tau) P(\mu \mid \tau) \qquad (2.2.13)$$

Here $P(\tau)$ represents when next reaction occurs and $P(\mu \mid \tau)$ reperents which reaction occurs. We need to generate two random numbers r_1 and r_2, which are uniformly distributed on $[0, 1]$. Then,

$$e^{-a_0\tau} = r_1 \quad \Rightarrow \quad \tau = (1/a_0)\ln(1/r_1) \tag{2.2.14}$$

and μ is the integer for which

$$\sum_{i=1}^{\mu-1} \frac{a_i}{a_0} < r_2 < \sum_{i=1}^{\mu} \frac{a_i}{a_0} \tag{2.2.15}$$

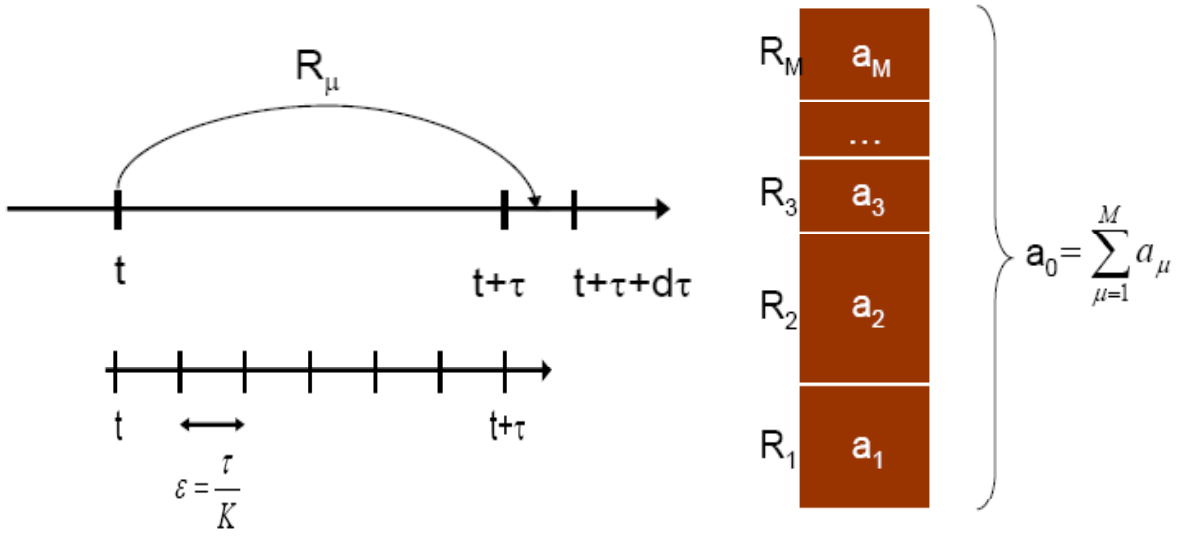

Figure 5: Scheme of the Gillespie's direct method [47]

Here is the pseudo algorithm of Gillespie's direct method:

1. Initialization:
   - Set initial numbers of molecules and time $t = 0$.
   - Set values of C_μ for the M reactions.
2. Calculate the values of a_i and $a_0 = \sum_{i=1}^{M} a_i$.
3. Generate the two uniformly distributed random numbers r_1 and r_2;

- Set $\tau = (1/a_0)\ln(1/r_1)$
- Set μ to the integer that satisfies: $\sum_{i=1}^{\mu-1} a_i/a_0 < r_2 < \sum_{i=1}^{\mu} a_i/a_0$.

4. Change the number of molecules to reflect the execution of reaction μ.
5. Increase time t by τ.
6. Return to step 2 until total simulation time t reaches target maximum time.

In each loop, DM generates two random number in equation (2.2.14) and (2.2.15). In the algorithm of FRM, it generates a τ_k for each reaction channel R_k according to $\tau_k = (1/a_k)\ln(1/r_k)$, where $k = 1, ..., M$, and r_1, $...r_M$ are statistically independent samplings from the unit interval. Then $\tau = min\{\tau_1, ...\tau_M\}$ and μ is the index of the $min\{\tau_1, ...\tau_M\}$. The improvement of NRM is that NRM reuses the $M-1$ unused randomly generated numbers in FRM and generates a new random number to replace the used one, then carefully updates the system to choose the next reaction. Clever data storage structures are employed to efficiently find τ and μ. The DM, FRM and NRM all produce exactly realizations of molecules by generating random pairs (τ, μ) rigorously according to the joint density function (2.2.13). The three methods can therefore be regarded as mathematically equivalent. NRM is widely believed to be more efficient than DM because it only generates one random number per step while DM generates two. But NRM has extra cost in its special data structures. Cao [10] tested the two algorithms on several systems and found NRM not as efficient as DM most of the time. Then they proposed a method called optimized direct method by re-indexing the reaction channels in a large system. The simulation time is speeded up by 12% - 25 % by Cao's result.

In our model, we choose the DM implementation of the SSA to simulate our system. All the algorithms are implemented in MATLAB 7.5.0 (R2007b). In order to speed up the simulation, we utilize the powerful feature of matrix computation in MATLAB and wrote the code in such a way that hundreds of independent trajectories can be simulated in a single loop. This increased simulation speed by a factor of $10 \sim 50$ compared with a program in which the trajectories are simulated consecutively. We have studied the system behavior for various values of n at several different system sizes. The system with $H^* = 100$ and with initial condition $(V_0, H_0, I_0) = (0.03H^*, H^*, 0)$ is the most common system size we have

studied because of its reasonable simulation time. We focus our attention on the following characteristics of the system:

Definition 2.2.1. Conditional mean extinction time, defined as the mean cell or virus extinction time over trajectories in which cell or virus extinction occurs.

Definition 2.2.2. Extinction probability, defined as the probability of cell or virus extinction.

Definition 2.2.3. Limiting extinction probability, defined as the limit of cell or virus extinction probability as $n \to \infty$.

Results: Figure 6 shows the sample trajectories of the system produced by Gillespie's direct method. The trajectories can be divided into three distinct types, depending on whether they terminate in healthy state (type I), dead state (type II), or they persist without termination (type III).

For small values of n, essentially all trajectories are of type I. At some intermediate values of n, roughly corresponding to the values of $\gamma_V^{BP} < \gamma_V < \gamma_V^{HB}$, almost all the trajectories are type III. At large values of n, most trajectories are of type II . Once a typical trajectory visits the neighborhood of the invariant facet $H = 0$ before visiting the neighborhood of the line $V = I = 0$, it is likely to be captured by the dead state.

In Figure 6, the initial number of healthy cells, viruses, and infected cells are $H_0 = H^* = 100$, $V_0 = 3$, and $I_0 = 0$. The other parameters are $n = \gamma_V/a_I = 15$ and maximum simulation time $T_{max} = 30$. The reaction rates used in the simulation are referred from Hancioglu's paper [40], $a_I = 1.5$, $a_V = 1.7$, $\gamma_{VH} = 0.7/H^*$, $b_{HD} = 4/H^*$ (Some of the parameters are adjusted by the homeostatic level in our Gillespie algorithm).

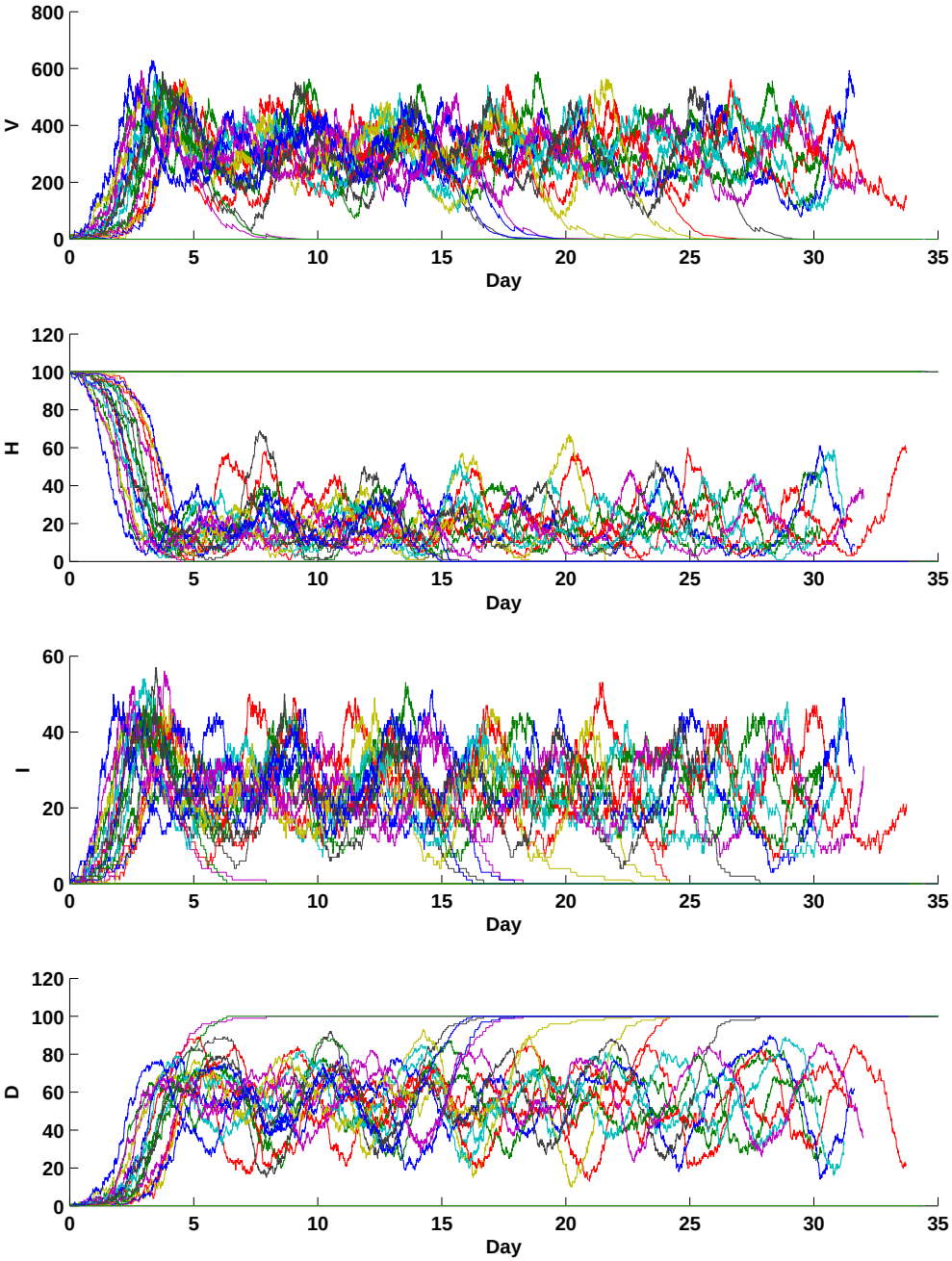

Figure 6: Sample trajectories of the basic system obtained by the Gillespie's direct method. $H_0 = H^* = 100$, $V_0 = 3$, $I_0 = 0$, $n = 15$, $T_{max} = 30$, number of the trajectories=30.

For comparison, we have run the ODE model and Gillespie simulation together by using the same parameter values (see Figure 7). We set $\gamma_V = 40$, so $n = \gamma_V/a_I = 27$, and

$T_{max} = 30$. The trajectories of ODE model are showed in the first row subplots. We can see the limit cycles in the $H - V$ and $I - V$ planes. Healthy cells H are always able to recover from an almost complete absence. The trajectories in the second row are derived from a single simulation using Gillespie's direct method. We can see that when healthy cells drop below a certain level, they try to recover, but after small excursion they go to extinction at $t \approx 6$. After healthy cell extinction, the infected cell is extinct at $t \approx 7$, and then followed the virus extinction at $t \approx 8$. The green dots in the plots are where the first extinction - healthy cell extinction in this sample - occurs.

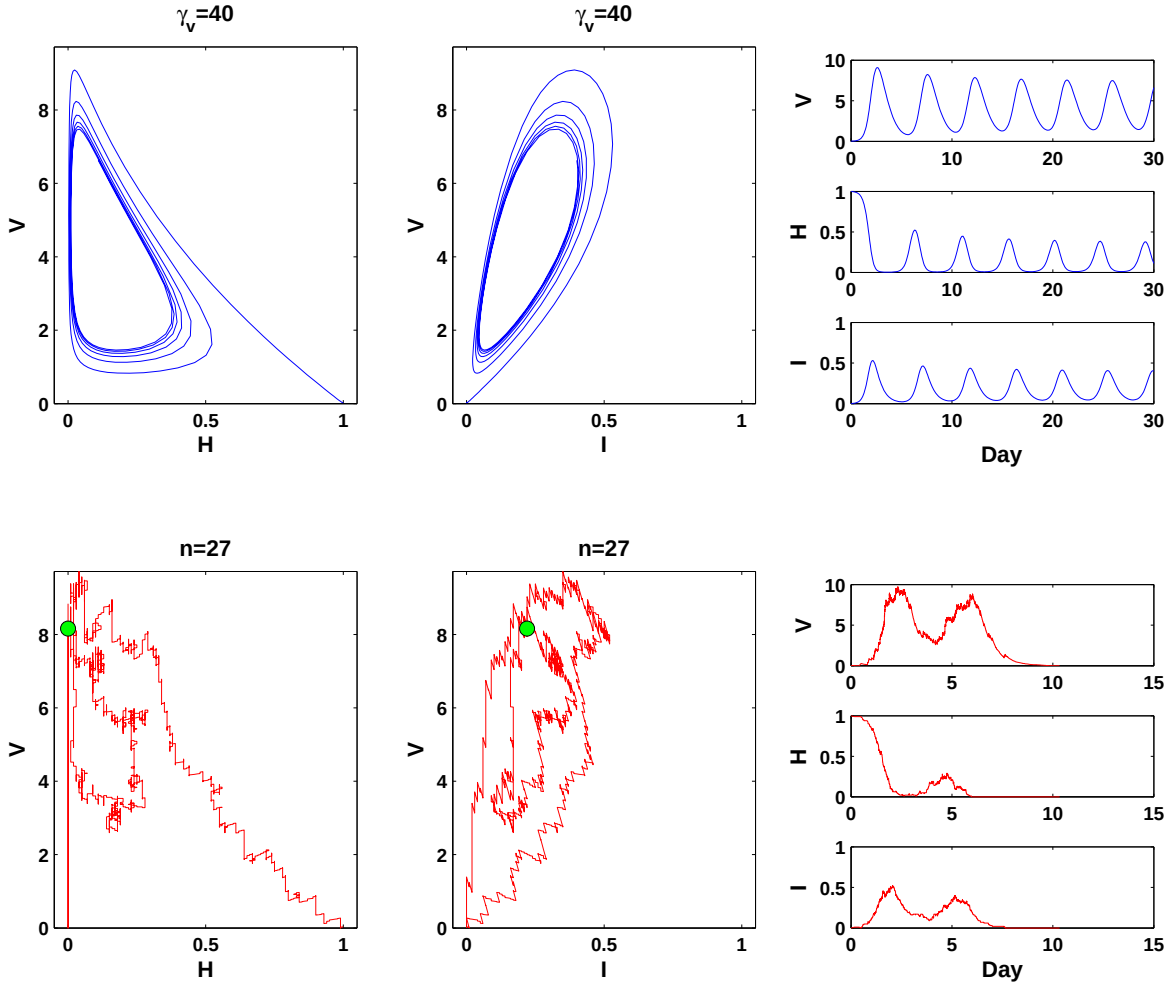

Figure 7: Trajectory comparison for ODE and stochastic models of the basic system. Parameters are equivalent in the two models. $(V_0, H_0, I_0) = (0.03, 1, 0)$ in the deterministic model and $(V_0, H_0, I_0) = (3, 100, 0)$ in the stochastic model. $\gamma_V = 40$, $n = 27$, $T_{max} = 30$.

In Figure 8, we plot the conditional mean extinction times and extinction probabilities as functions of n. As we mentioned before, the stochastic model shows three kinds of behavior in the different parameter regions. When $n < 6$, no cell extinction occurs. When $6 \leq n < 11$, a few cell extinctions occurs and the conditional mean time of cell extinction increases dramatically. When $11 \leq n < 25$, the cell extinction probability increases fast to a limiting probability $P_C = 0.64$, and remains at that level as $n \to \infty$. At the same time, the conditional mean cell extinction time gradually decreases to a certain level. For the viruses, it has an almost complete extinction with probability 1 when n is small, decreases fast when $n < 10$, and gradually decreases to a limiting probability $P_V = 0.36$ when $n > 10$. The conditional mean virus extinction time has a peak at around $n = 5$, where the cell extinction starts to appear, and then it decreases fast at $5 \leq n < 10$ and reach a limiting extinction time when $n > 25$.

We have examined the effect of simulation time on the results when $T_{max} = 10, 20, 30, 40, 50, 60$, and define the **non-extinction** as neither cell nor virus extinction occurs when the simulation time reaches T_{max}. In Figure 8, the trajectories of conditional extinction time and extinction probability at different T_{max} are the same before the bifurcation and at the limiting case. The larger T_{max}, the faster the system reaches the limiting or stable level. As expected, the peaks of the cell and virus conditional extinction time always occur at the same n intervals and become higher as we increase the simulation time. According to Figure 8, it appears that there might be an explosion of the extinction time at some certain values of n. Below we verify that such an explosion does not happen, i.e., the mean extinction time is finite, using the Laplace transform method.

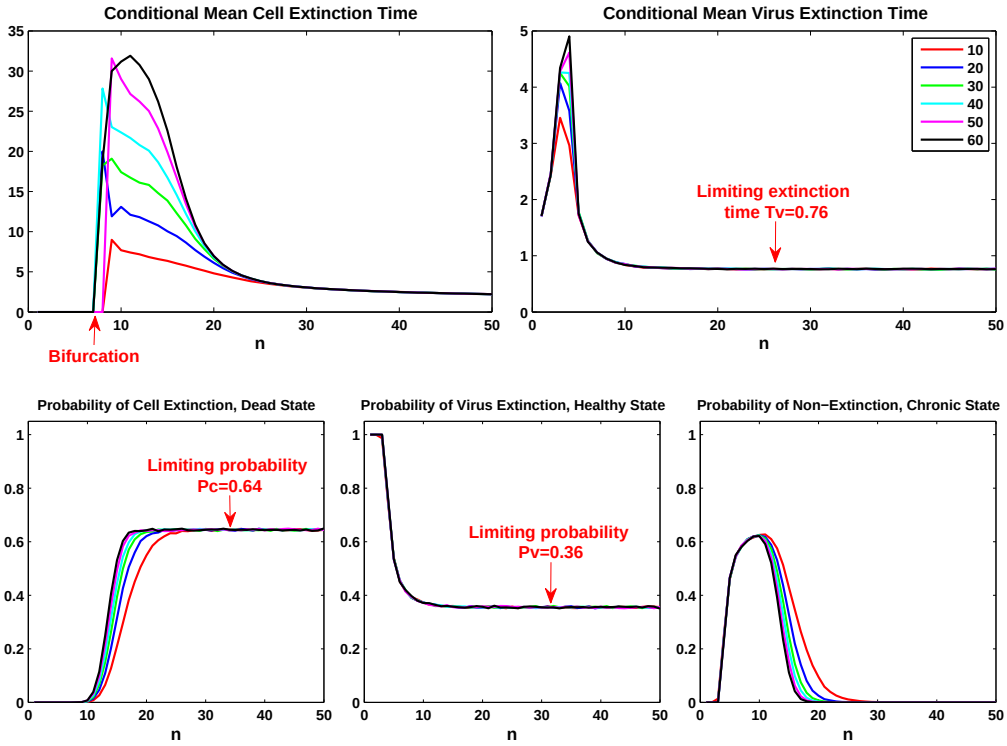

Figure 8: Conditional mean time and probability of extinction with different maximum simulation times, $T_{max} = 10, 20, 30, 40, 50, 60$, $H^* = 100$, $V_0 = 3$. The x-axis represents n values and the y-axis represents conditional mean extinction time in the first row plots and extinction probability in the second row plots. All the trajectories are averaged over 50,000 independent samples.

2.2.2.2 Numerical Solution of Mean Extinction Time The master equation for the system (2.2.7)-(2.2.10) was given in (2.2.11). It is too complex to be solved analytically, but it can be solved numerically if the state space is truncated at an upper limit of the copy number in each species. In fact, since the total number of cells is limited by H^*, one only needs to truncate the number of virus particles. An artificial boundary condition $P(x,t) = 0$ is usually added at $x_i > x_{max}$ for a sufficiently large $x_{max} > 0$ to obtain a finite computational domain.

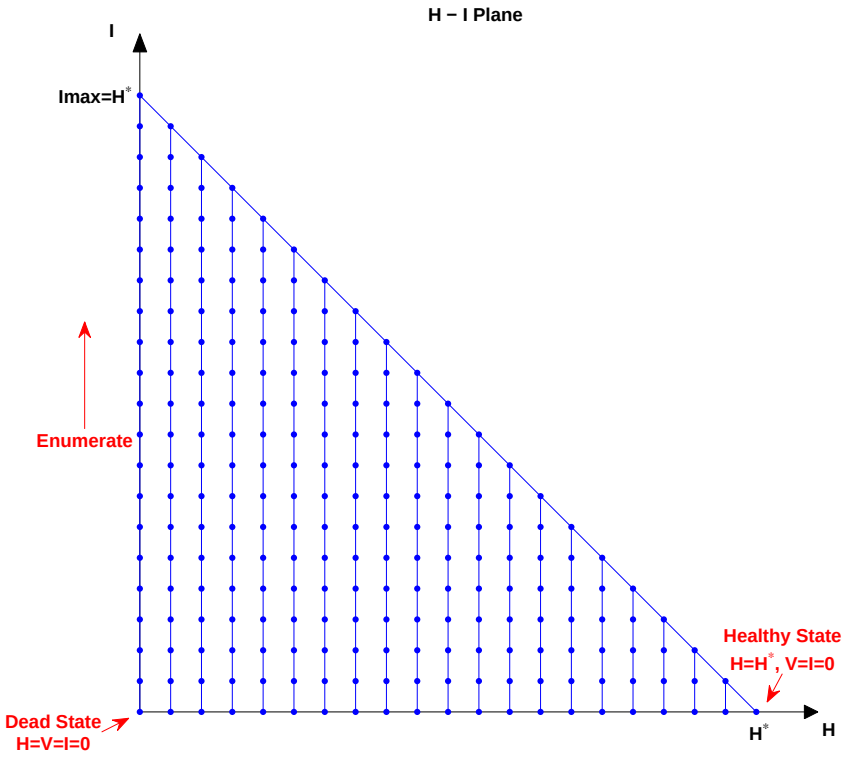

Figure 9: Enumerating the states on $H - I$ plane

We truncate the state space at an upper limit V_{max} of the copy number of viruses. Variables H, I, D are bounded by the equation $H + I + D = H^*$. The remaining problem would be how to enumerate all the states in the truncated system. Figure 9 shows our method to enumerate the states on $H - I$ plane. We are going to enumerate all the states $H - I$ plane by $H - I$ plane according to different number of viruses. We start at the $H - I$ plane with $V = 0$. The dead state $(V, H, I) = (0, 0, 0)$ is the first state, $(V, H, I) = (0, 0, 1)$ is the second state, then $(V, H, I) = (0, 0, 2)$ is the third state and so on. The healthy state $(V, H, I) = (0, H^*, 0)$ is the last state on the first plane $(V = 0)$. Then we continue enumerating on the second $H - I$ plane with $V = 1$, and then the third plane with $V = 2$ till $V = V_{max}$. The total number of states is:

$$N = (H^* + 1)(H^* + 2)(V_{max} + 1)/2 \tag{2.2.16}$$

If we take $H^* = 100$, $V_{max} = 600$, the system size is approximately 3×10^6. It is still so large that it is not possible to solve the master equation using numerical ODE solvers. Fortunately, there is a way to compute exactly the mean extinction time and other moments by using Laplace transform of the master equation [67].

Here is how we implement the method for our model. First, each P_i corresponds to a $P_{V,H,I}$ with V, H, I being a certain number. P_1, P_2, P_3, ..., $P_{TotalPoint}$ denote the probability at each state. P_H, P_D are the total probabilities of states in healthy absorbing region ($V = I = 0$, $H > 0$, red line in Figure 4) and dead absorbing region ($H = 0$, green plane in Figure 4). So we have:

$$P_H = \sum_{H=1}^{H^*} P_{0,H,0} \tag{2.2.17}$$

$$P_D = \sum_{V=0}^{V_{max}} \sum_{I=0}^{H^*} P_{V,0,I} \tag{2.2.18}$$

The probability for the system to go to healthy absorbing states in a small time interval δ_t is $P_H(t+\delta_t) - P_H(t)$. Shrinking this time interval to zero, we have the distribution of the extinction time:

$$\frac{dP_H}{dt} = \lim_{\delta_t \to 0} \frac{P_H(t+\delta_t) - P_H(t)}{\delta_t} \tag{2.2.19}$$

From (2.2.11), we know

$$\begin{aligned} \sum_{H=1}^{H^*} \frac{dP_{0,H,0}}{dt} &= \sum_{H=1}^{H^*} \Big(a_V P_{1,H,0}(t) + b_{HD}(H-1)(H^*-H+1) P_{0,H-1,0}(t) \\ &\quad -(b_{HD} H(H^*-H)) P_{0,H,0}(t) \Big) \qquad (2.2.20) \\ &= \sum_{H=1}^{H^*} a_V P_{1,H,0}(t) \qquad (2.2.21) \end{aligned}$$

Thus the distribution of the virus extinction is simply:

$$\frac{dP_H}{dt} = \sum_{H=1}^{H^*} a_V P_{1,H,0}(t) \tag{2.2.22}$$

So the moments of the virus extinction time are obtained via

$$\langle {T_V}^n \rangle = a_v \int_0^\infty t^n \Big(\sum_{H=1}^{H^*} P_{1,H,0}(t) \Big) dt \tag{2.2.23}$$

Similarly, the distribution of the cell extinction is

$$\frac{dP_D}{dt} = \sum_{V=1}^{V_{max}} \sum_{I=0}^{H^*-1} \gamma_{VH} V P_{V,1,I}(t) \tag{2.2.24}$$

The moments of the cell extinction time are obtained via

$$\langle {T_C}^n \rangle = \gamma_{VH} \int_0^\infty t^n \Big(\sum_{V=1}^{V_{max}} \sum_{I=0}^{H^*-1} V P_{V,1,I}(t) \Big) dt \tag{2.2.25}$$

The Laplace transform of function $P(t)$ is defined via

$$\widehat{P}(s) = \int_0^\infty e^{-st} P(t) dt \tag{2.2.26}$$

and the transform of the derivative $P'(t)$ by

$$\int_{0^-}^\infty e^{-st} P'(t) dt = s \int_{0^-}^\infty e^{-st} P(t) dt - P(0) \tag{2.2.27}$$

The derivative of $\widehat{P}(s)$ with respect to s can be related to $P(t)$ by

$$\frac{d^n \widehat{P}(s)}{ds^n} = \int_0^\infty (-t)^n e^{-st} P(t) dt \tag{2.2.28}$$

Thus, can rewrite the moments equations in the following way:

$$\langle {T_V}^n \rangle = (-1)^n a_V \sum_{H=1}^{H^*} \frac{d^n \widehat{P}_{1,H,0}(s)}{ds^n} \Big|_{s=0} \tag{2.2.29}$$

$$\langle {T_C}^n \rangle = (-1)^n \gamma_{VH} \sum_{V=1}^{V_{max}} \sum_{I=0}^{H^*-1} V \frac{d^n \widehat{P}_{V,1,I}(s)}{ds^n} \Big|_{s=0} \tag{2.2.30}$$

Now we are almost ready to calculate the moments of extinction times. The difficulty of solving the master equation is the huge system size of (2.2.11). Our plan is to transform the master equation (2.2.11) into a linear system through its Laplace transform, then solve the linear system, and substitute the solutions in to equations (2.2.29) and (2.2.30) to get the values of moments generating function. Apply (2.2.26) and (2.2.27) to the right and left

sides of (2.2.11) separately. The Laplace transform of the master equation (2.2.11) is the following system of algebraic equations:

$$\begin{aligned} s\widehat{P}_{V,H,I}(s) - P_{V,H,I}(0) = \gamma_{VH}(H+1)(V+1)\widehat{P}_{V+1,H+1,I-1}(s) + a_I(I+1)\widehat{P}_{V-n,H,I+1}(s) \\ +a_V(V+1)\widehat{P}_{V+1,H,I}(s) + b_{HD}(H-1)(H^*-H-I+1)\widehat{P}_{V,H-1,I}(s) \\ -(\gamma_{VH}HV + a_I I + a_V V + b_{HD}H(H^*-H-I))\widehat{P}_{V,H,I}(s) \end{aligned} \tag{2.2.31}$$

Equation (2.2.31) is a linear system for any value of s. Let's set $s = 0$ and denote it as the following.

$$B_0 = A\widehat{P}_{V,H,I}(0) \tag{2.2.32}$$

The initial condition of (2.2.32) is $P_{V_0,H^*,0}(0) = 1$ and $P_{V,H,I}(0) = 0$ for all other states. $B_0 = -P_{V,H,I}(0)$ and A is the coefficient matrix of the right hand side of the system(2.2.31). Equations (2.2.29) and (2.2.30) tell us that the moments of the extinction time depend on the derivatives of the $\widehat{P}_{V,H,I}$ functions. By taking the derivative of (2.2.31) with respect to s we obtain

$$\begin{aligned} \widehat{P}_{V,H,I}(s) + s\frac{d\widehat{P}_{V,H,I}(s)}{ds} = \gamma_{VH}(H+1)(V+1)\frac{d\widehat{P}_{V+1,H+1,I-1}(s)}{ds} + a_I(I+1)\frac{d\widehat{P}_{V-n,H,I+1}(s)}{ds} \\ +a_V(V+1)\frac{d\widehat{P}_{V+1,H,I}(s)}{ds} + b_{HD}(H-1)(H^*-H-I+1)\frac{d\widehat{P}_{V,H-1,I}(s)}{ds} \\ -(\gamma_{VH}HV + a_I I + a_V V + b_{HD}H(H^*-H-I))\frac{d\widehat{P}_{V,H,I}(s)}{ds} \end{aligned} \tag{2.2.33}$$

Set $s = 0$, equation (2.2.33) becomes a linear system again. Let us denote it as

$$B_1 = A\frac{d\widehat{P}_{V,H,I}(0)}{ds} \tag{2.2.34}$$

Where $B_1 = \widehat{P}_{V,H,I}(0)$, which can be solved from the linear system (2.2.32). By repeating the above process one can obtain higher moments.

The results can be summarized in the following algorithm:

1. Solve the linear system (2.2.32) to get $\widehat{P}_{V,H,I}(0)$.
2. Substitute the solutions of (2.2.32) into (2.2.34), and solve the system to get $\frac{d\widehat{P}_{V,H,I}(s)}{ds}$.
3. Calculate the mean extinction time of virus through formula (2.2.29) by setting $n = 1$,

$$\langle T_V \rangle = -a_V \sum_{H=1}^{H^*} \frac{d\widehat{P}_{1,H,0}(0)}{ds}\Big|_{s=0} \tag{2.2.35}$$

4. Calculate the mean extinction time of cell through formula (2.2.30) by setting $n = 1$,

$$\langle T_C \rangle = -\gamma_{VH} \sum_{V=1}^{V_{max}} \sum_{I=0}^{H^*-1} V \frac{d\widehat{P}_{V,1,I}(s)}{ds}\Big|_{s=0} \tag{2.2.36}$$

To calculate the second moments, take the second derivative of equation (2.2.31) and set $s = 0$, we get

$$B_2 = A\frac{d^2\widehat{P}_{V,H,I}(0)}{ds^2} \tag{2.2.37}$$

Where $B_2 = 2\frac{d\widehat{P}_{V,H,I}(0)}{ds}$, which is solved from the linear system (2.2.34). Similarly, solving the linear system (2.2.37) to get $\frac{d^2\widehat{P}_{V,H,I}(0)}{ds^2}$ and substitute the solution back to formula (2.2.29) and (2.2.30) by setting $n = 2$ to get the second moments of the extinction times. The variance of the extinction times is given by

$$Var_V = \langle {T_V}^2 \rangle - \langle T_V \rangle^2 \tag{2.2.38}$$

$$Var_C = \langle {T_C}^2 \rangle - \langle T_C \rangle^2 \tag{2.2.39}$$

and the coefficient of variation is given by

$$\rho_V = \frac{\sqrt{\langle {T_V}^2 \rangle - \langle T_V \rangle^2}}{\langle T_V \rangle} \tag{2.2.40}$$

$$\rho_C = \frac{\sqrt{\langle {T_C}^2 \rangle - \langle T_C \rangle^2}}{\langle T_C \rangle} \tag{2.2.41}$$

To calculate the higher moments of extinction, we just need to continue taking derivatives of (2.2.31) and solving the new linear system by substituting solutions from the previous step.

We have mentioned before that the main advantage of the Laplace transform method is to convert the nonlinear master equation into a linear system of algebraic equations. Unfortunately, if we take $H^* = 100$, $V_{max} = 600$, the system size is approximately 3×10^6.

Because of limited computer memory, we were not able to handle a system size as large as 3×10^6. We have explored several ways to improve the computational efficiency. Suppose the system that we are going to solve is $\bar{B} = \bar{A}\bar{x}$. First, $\bar{A}$ is a sparse matrix in our model. We reorder the $\bar{A}$ by Cuthill-McKee ordering [15] - a permutation such that $\bar{A}$ tends to have its nonzero elements closer to the diagonal. This reordering is very helpful in LU factorization. We then use $B = Ax$ to denote the new system. We divide the matrix A into 4 sub-matrices with the same size, B into 2 vectors with same length and so does x. Then we have:

$$\begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{2.2.42}$$

That is

$$B_1 = A_{11}x_1 + A_{12}x_2 \tag{2.2.43}$$

$$B_2 = A_{21}x_1 + A_{22}x_2 \tag{2.2.44}$$

One can solve for x_1 using

$$(A_{12}^{-1}A_{11} - A_{22}^{-1}A_{21})x_1 = (A_{12}^{-1}B_1 - A_{22}^{-1}B_2) \tag{2.2.45}$$

then substitute x_1 into (2.2.43) and solve for x_2. Each of the matrices in question has a lower size than the original matrix A and hence they all fit in the computer memory. We apply LU factorization whenever necessary into the intermediate matrices in above procedure.

This method help in improving the upper limit on the size of a solvable linear system in MATLAB. It requires less memory when dealing with the inverse of these sub-matrices than a full matrix. In the end, the largest size of the system that we were capable to solve was $H^* = 25$, $V_{max} = 181$, corresponding to a total of 63882 states in the system.

Figure 10 shows the result of Laplace method. We have also applied the same parameter set to Gillespie algorithm and plot the results of these two methods together. Notice that equation (2.2.29) and (2.2.30) are formulas of mean extinction times. But what we have calculated in the Gillespie's direct method is the conditional mean extinction times. So we need to apply the probability values at each n values derived from Gillespis algorithm to (2.2.29) and (2.2.30) to get the conditional mean extinction time.

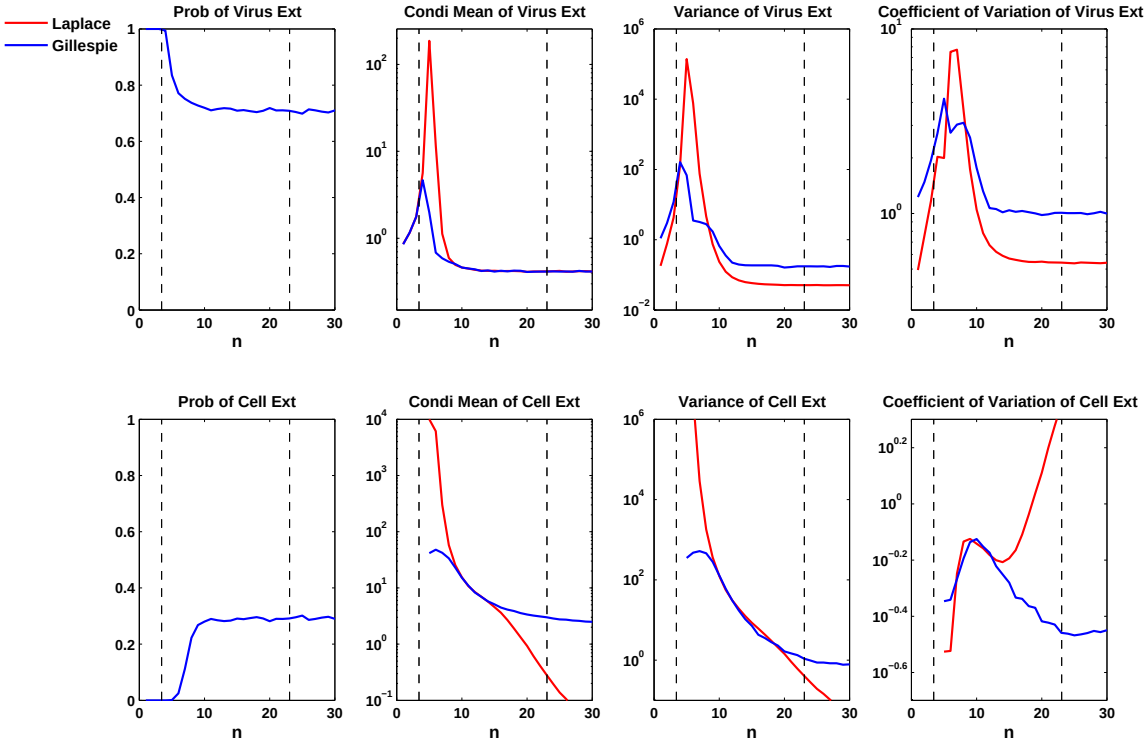

Figure 10: Probability, conditional mean, variance and coefficient of variation of virus or cell extinction. The red line is the result of Laplace method and the blue line is the result of Gillespie's direct method. $V_0 = 1$, $H^* = 25$, $V_{max} = 181$, all the other parameters are the same as before. The x-axis is $n = \gamma_V / a_I$.

This Laplace method only gives us the moments of the extinction times, not extinction probabilities. The two plots in the first column in Figure 10 are the probabilities of virus and cell extinction computed by the Gillespie's method. The plots in the second column are the conditional mean extinction times of the virus and cell. As we have pointed out before, because of the limited simulation time, the peak in the Gillespie's method is not accurate. The plots in the third and forth column are the variance and coefficient of variation of virus and cell. The y-axis of the plots are in log scale except ones in the first column. We see that for $n < 10$, Laplace method generates much higher peaks than the Gillespie's direct method. When $n > 10$, the two methods coincide for the conditional mean extinction time

of virus and have similar trajectories for the variance and coefficient of variation of virus extinction. Because cell extinction is very sensitive to V_{max}, the results of the two methods for cell extinction time are similar only at some intermediate interval of n values.

From the computational side, Laplace method is faster than Gillespie's direct method. But the Gillespie's method works better in the case of cell extinction than Laplace method does, especially at high values of n where trajectories make excursions to high values of V. When increased the system size, the variance of cell extinction in Laplace method became negative at some values of n which implies that the system size has become to large to be accurately computed by Laplace method. So Laplace method is accurate and efficient when H^* and n are small.

In summary, Gillespie's direct method is limited by the simulation time and Laplace method is limited by the system size, especially the cutoff value of V_{max}. In order to take the advantage of both methods, we combine the two methods to apply at different values of H^* and n.

2.2.2.3 Combination of Gillespie and Laplace Method In last section, we mentioned that the Laplace method is more efficient and accurate at small system size and when n is small while Gillespie's method is more accurate for large n. Figure 11 shows the result of the combination of the two methods.

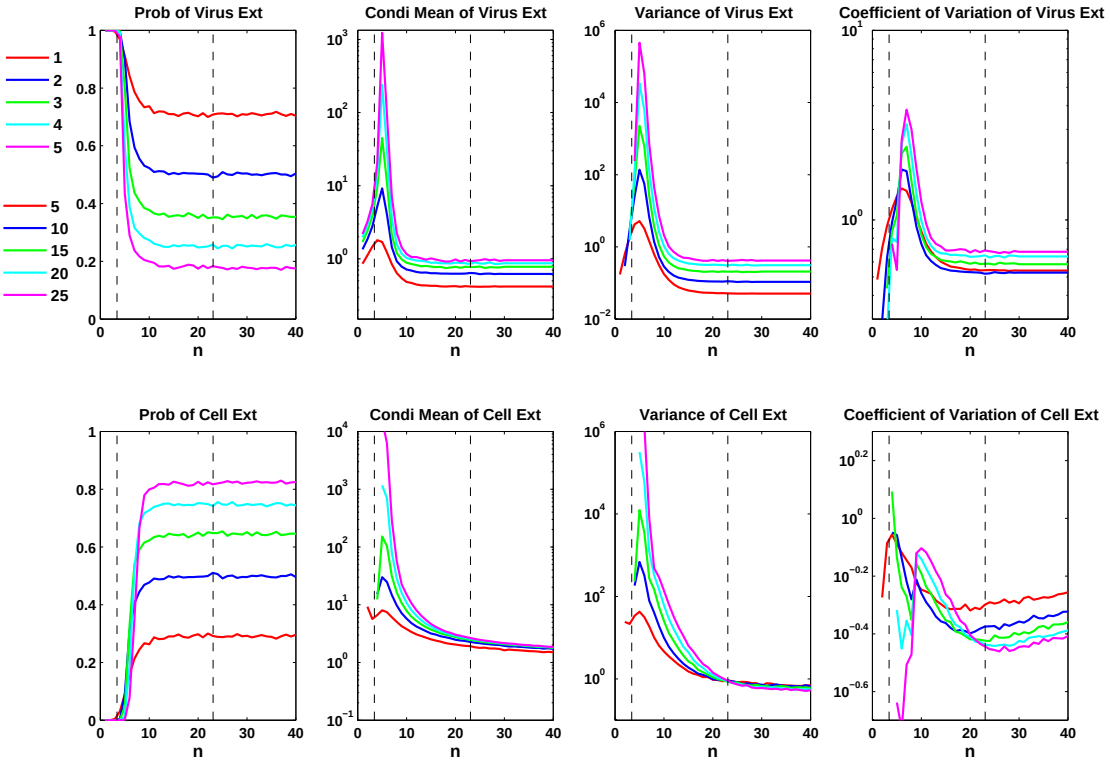

Figure 11: A combination of Gillespie's direct and Laplace methods. The red, dark blue, green, light blue and pink trajectories represent $(H^*, V_0) = (5, 1), (10, 2), (15, 3), (20, 4)$, and $(25, 5)$ with $V_{max} = 181$. The x-axis is $n = \gamma_V / a_I$. Note the logarithmic scale of the ordinate in the plots of extinction times, variances and coefficients of variation.

We have studied the dependence of extinction times on system size by using the following values, $(H^*, V_0) = (5, 1), (10, 2), (15, 3), (20, 4), (25, 5)$. V_{max} was fixed at 181. Figure 11 shows how the system size effect the results. The limiting probability of virus extinction decreases as we increase the system size while the limiting probability of cell extinction increases.

In both Gillespie's and Laplace's methods, we find that the conditional extinction times of both virus and cells grow exponentially with the system size. One dimensional models studies by Nasell [66] and Vellela [83] agree with this result. In our model, the coefficient of variation of virus extinction increases with the system sizes and reaches the maximum value

at some intermediated n values. Newman [67] states in a one dimensional birth-death model that the coefficient of variation is found to have a maximum value when the death and birth rates are close in the value. These paper all discuss the extinction time in three parameter regions: long, intermediate long and short extinction time. The second region is a transition region between the other two.

2.2.2.4 Approximate Solution of a Limiting Case Recall from Figure 6 that there is a special subset of trajectories which are of short duration and in which the number of viruses never increases. In this section, we will show that these trajectories have the same probability of occurrence independent of n and can be interpreted as *no-disease trajectories*. We analyze them analytically using a corresponding Markov Chain.

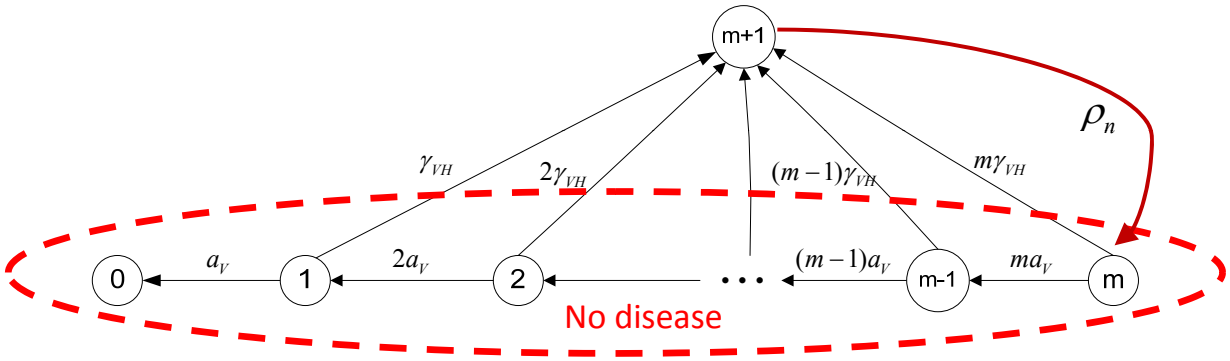

Figure 12: m is the initial number of viruses. 0, 1, 2, ..., m are the states representing $V = 0, 1, ...m$ and $I = 0$. $m + 1$ represents all states with $V > m$ or $I > 0$. The red dashed ellipse denotes the no-disease region.

In Figure 12, numbers 0, 1, 2, ..., m represent the states $(V, H, I) = (0, H^*, 0), ..., (m, H^*, 0)$ of the system. Thus, the state 0 represents the healthy state. The state $m + 1$ is outside of the no-disease region with $I > 0$ or $V > m$ (assume $m < n$). The only way for the system to leave the no-disease region and enter state $m + 1$ is when reaction $V + H \to I$ occurs. The only channel for the system to go from state $m + 1$ back to no-disease region is that it goes to state m first. Let's denote this returning probability as ρ_n. State m has $I = 0$. So $I \to nV + D$ must occur with probability 1 if the system is to go from $m + 1$ back to state

m. Once $I \to nV + D$ occurs, the system has $\geq n$ number of viruses, the probability for these viruses to decay to number m is $P_{m+1-to-m} = \left(\frac{a_V}{a_V+\gamma_{VH}}\right)^{n-m}$. So $\rho_n \leq P_{m+1-to-m}$. As n becomes large, we have:

$$0 \leq \lim_{n\to\infty} \rho_n \leq \lim_{n\to\infty} \left(\frac{a_V}{a_V + \gamma_{VH}}\right)^{n-m} = 0 \tag{2.2.46}$$

Therefore, as $n \to \infty$, and $\lim_{n\to\infty} \rho_n = 0$, the basic system behaves as the Markov chain in Figure 12. In the other word, as $n \to \infty$ the basic system will either stay in the no -disease region, and converge to the healthy state, or it will leave the no-disease region and never return to healthy state, as $m + 1$ becomes an absorbing state. The master equation of this reduced model is below:

$$\begin{aligned}
\frac{dP_{m+1}}{dt} &= \gamma_{VH}P_1 + 2\gamma_{VH}P_2 + 3\gamma_{VH}P_3 + \cdots + m\gamma_{VH}P_m \\
\frac{dP_0}{dt} &= a_V P_1 \\
\frac{dP_1}{dt} &= -(a_V + \gamma_{VH})P_1 + 2a_V P_2 \\
\frac{dP_2}{dt} &= -2(a_V + \gamma_{VH})P_2 + 3a_V P_3 \\
&\vdots \\
\frac{dP_{m-1}}{dt} &= -(m-1)(a_V + \gamma_{VH})P_{m-1} + ma_V P_m \\
\frac{dP_m}{dt} &= -m(a_v + \gamma_{VH})P_m
\end{aligned} \tag{2.2.47}$$

We have employed two methods to solve the system (2.2.47).

Method 1: The Embedded Markov Chain

If the states of a Markov process are discrete and the process may change state at any point in time, we say that the process is a continuous-time Markov chain (**CTMC**). If the transitions between states is independent of time, it is called a homogeneous, continuous-time Markov chain. Let $\bar{q}_{ij}$ be the rate at which transitions occur from stat i to state j. When the CTMC is homogeneous, the transition rates $\bar{q}_{ij}$ are independent of time and the matrix of transition rates can be written as $\bar{Q}$. When discussing the Gillespie algorithm, we have shown that the time spent by the system in any state is exponentially distributed.

In a state i, the parameter of the exponential distribution of the sojourn time is given by $-\bar{q}_{ii} = \sum_{j \neq i} \bar{q}_{ij}$. If we ignore the time actually spent in any state and consider only the transitions made by the system, we may define a new, discrete-time, Markov chain called the *embedded Markov chain* (**EMC**). Many of the properties of a CTMC can be deduced from its corresponding EMC. For example, a CTMC is irreducible if and only if its EMC is irreducible and a single state is recurrent in a CTMC if and only if it is recurrent in its EMC [78].

For a CTMC the one step transition probabilities of its corresponding EMC, denoted by $q_{i,j}$, are given by

$$\begin{aligned} q_{ij} &= \frac{\bar{q}_{ij}}{\sum_{j \neq i} \bar{q}_{ij}}, \quad j \neq i \\ &= 0, \quad i = j \end{aligned} \tag{2.2.48}$$

Let Q denote the transition probability matrix of the EMC. All elements in Q satisfy $0 \leq q_{ij} \leq 1$ and $\sum_{i,j \neq i} q_{ij} = 1$ for all i. Thus Q possesses the characteristics of a transition probability matrix for a discrete-time Markov chain [78]. So the limiting probability can be found as

$$P = \lim_{n \to \infty} P_0 Q^n \tag{2.2.49}$$

where P_0 is the initial probability distribution.

We shall use the EMC to find the limiting probability of virus extinction. Let's start with examples in which the initial number of viruses is 1 and 2.

Case 1: $m = 1$

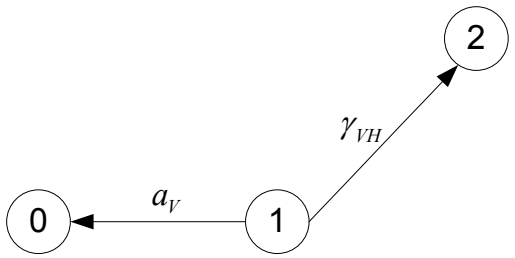

Figure 13: State 1 is the initial state $(V, H, I) = (1, H^*, 0)$, state 0 is the healthy state $(V, H, I) = (0, H^*, 0)$, and state 2 is the dead absorbing region $I > 0$ or $V > 1$.

State 2 represents the absorbing region outside of the no-disease region. State 0 is the healthy state with $(V, H, I) = (0, H^*, 0)$. State 1 is the initial state with $(V, H, I) = (1, H^*, 0)$. We shall reorder the states as $(2, 0, 1)$ so that we can get the transition rate matrix in a good permutation for computation, i.e., $P = (P_2, P_0, P_1)$. The transition probability matrix is:

$$Q_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{\gamma_{VH}}{a_V+\gamma_{VH}} & \frac{a_V}{a_V+\gamma_{VH}} & 0 \end{pmatrix} \tag{2.2.50}$$

and the initial state $P(0) = (0, 0, 1)$. If we take Q_1 squared, we find:

$$Q_1^2 = \frac{1}{(a_V+\gamma_{VH})^2} \begin{pmatrix} a_V+\gamma_{VH} & 0 & 0 \\ 0 & a_V+\gamma_{VH} & 0 \\ \gamma_{VH} & a_V & 0 \end{pmatrix}^2 = Q_1 \tag{2.2.51}$$

Hence

$$Q_1^l = Q_1^{l-1} = \ldots = Q_1^2 = Q_1, \quad l \geq 1 \tag{2.2.52}$$

So

$$P(0)Q_1^l = (0, 0, 1)Q_1 = \left(1 - \frac{a_V}{a_V+\gamma_{VH}}, \frac{a_V}{a_V+\gamma_{VH}}, 0\right), \quad l \geq 1 \tag{2.2.53}$$

is the stationary state, which means that by the end the system has a probability of $1-\frac{a_V}{a_V+\gamma_{VH}}$ to be a dead state and a probability of $\frac{a_V}{a_V+\gamma_{VH}}$ to be a healthy state if the initial number of viruses is 1.

Case 2: $m = 2$

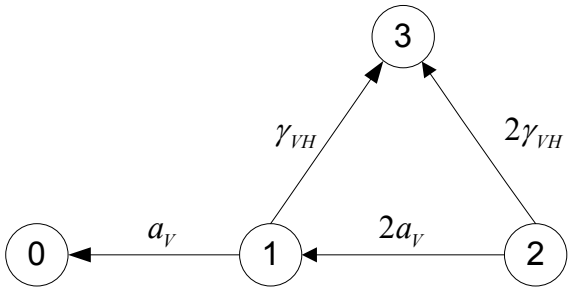

Figure 14: State 2 is the initial state, state 0 is the healthy state, and state 3 is the dead absorbing region.

Let $P = (P_3, P_0, P_1, P_2)$. The transition probability matrix is:

$$Q_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{\gamma_{VH}}{a_V+\gamma_{VH}} & \frac{a_V}{a_V+\gamma_{VH}} & 0 & 0 \\ \frac{\gamma_{VH}}{a_V+\gamma_{VH}} & 0 & \frac{a_V}{a_V+\gamma_{VH}} & 0 \end{pmatrix} \tag{2.2.54}$$

and the initial state $P(0) = (0, 0, 0, 1)$.

Calculate Q_2 square,

$$Q_2^2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1-\frac{a_V}{a_V+\gamma_{VH}} & \frac{a_V}{a_V+\gamma_{VH}} & 0 & 0 \\ 1-\left(\frac{a_V}{a_V+\gamma_{VH}}\right)^2 & \left(\frac{a_V}{a_V+\gamma_{VH}}\right)^2 & 0 & 0 \end{pmatrix} \tag{2.2.55}$$

Calculation of Q_2^3, reveals that $Q_2^3 = Q_2^2$. Hence,

$$Q_2^l = Q_2^{l-1} = \ldots = Q_2^2, \quad l \geq 2 \tag{2.2.56}$$

So

$$P(0)Q_2^l = (0,0,0,1)Q_2^2 = \left(1 - \frac{a_V^2}{(a_V+\gamma_{VH})^2}, \frac{a_V^2}{(a_V+\gamma_{VH})^2}, 0\right), \quad l \geq 2 \tag{2.2.57}$$

is the stationary state, which means that by the end the system has a probability of $1 - \left(\frac{a_V}{a_V+\gamma_{VH}}\right)^2$ to be a dead state and a probability of $\left(\frac{a_V}{a_V+\gamma_{VH}}\right)^2$ to be a healthy state if the initial number of viruses is 2.

On the basis of these two examples, we can deduce a general result.

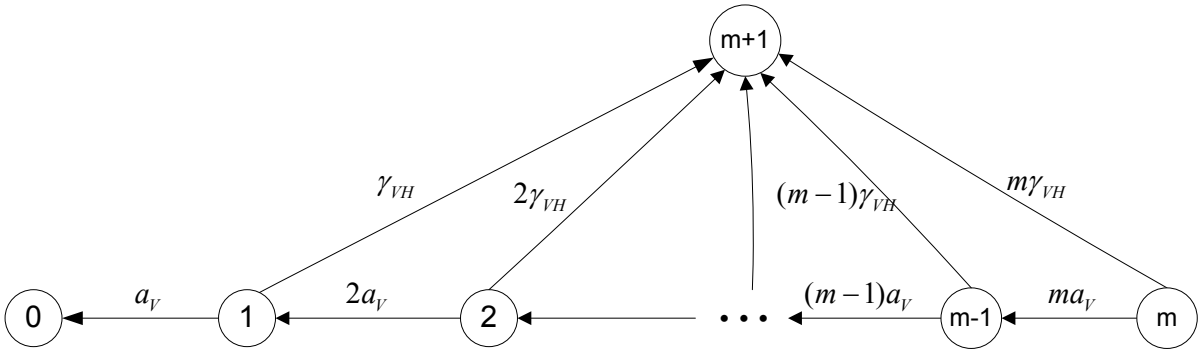

Figure 15: m is the initial number of viruses. 0, 1, 2, ..., m are the states representing the number of viruses and $I = 0$. $m+1$ represents the states when $V > m$ or $I > 0$.

Theorem 2.2.1. *If the initial number of viruses is m and $V + H \xrightarrow{\gamma_{VH}} I$ occurs, specifically when $n \to \infty$, the system (2.2.47) has a stationary state with the probability of virus extinction*

$$P_H = \left(\frac{a_V}{a_V+\gamma_{VH}}\right)^m \tag{2.2.58}$$

and the probability of cell extinction

$$P_D = 1 - \left(\frac{a_V}{a_V+\gamma_{VH}}\right)^m \tag{2.2.59}$$

The transition probability matrix Q_m has the property

$$Q_m^l = Q_m^{l-1} = \ldots = Q_m^{m+1} = Q_m^m, \quad l \geq m \tag{2.2.60}$$

where

$$Q_m^m = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 1-\frac{a_V}{a_V+\gamma_{VH}} & \frac{a_V}{a_V+\gamma_{VH}} & 0 & \cdots & 0 & 0 \\ 1-\left(\frac{a_V}{a_V+\gamma_{VH}}\right)^2 & \left(\frac{a_V}{a_V+\gamma_{VH}}\right)^2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & 0 \\ 1-\left(\frac{a_V}{a_V+\gamma_{VH}}\right)^{m-1} & \left(\frac{a_V}{a_V+\gamma_{VH}}\right)^{m-1} & 0 & \cdots & 0 & 0 \\ 1-\left(\frac{a_V}{a_V+\gamma_{VH}}\right)^{m} & \left(\frac{a_V}{a_V+\gamma_{VH}}\right)^{m} & 0 & \cdots & 0 & 0 \end{pmatrix} \tag{2.2.61}$$

Proof. We have verified this result for $m = 1, 2$. For $m \geq 3$, we prove the result using induction. Suppose it is true when $V_0 = m-1$, let's consider the case of $V_0 = m$. The transition probability matrix Q_m is a $(m+2) \times (m+2)$ matrix:

$$Q_m = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \frac{\gamma_{VH}}{a_V+\gamma_{VH}} & \frac{a_V}{a_V+\gamma_{VH}} & 0 & \cdots & 0 & 0 & 0 \\ \frac{\gamma_{VH}}{a_V+\gamma_{VH}} & 0 & \frac{a_V}{a_V+\gamma_{VH}} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & 0 & 0 \\ \frac{\gamma_{VH}}{a_V+\gamma_{VH}} & 0 & 0 & \cdots & \frac{a_V}{a_V+\gamma_{VH}} & 0 & 0 \\ \frac{\gamma_{VH}}{a_V+\gamma_{VH}} & 0 & 0 & \cdots & 0 & \frac{a_V}{a_V+\gamma_{VH}} & 0 \end{pmatrix} \tag{2.2.62}$$

Divide matrix Q_m into four sub-matrices,

$$Q_m = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{2.2.63}$$

where A is a $(m+1) \times (m+1)$ matrix composed by the intersection of the first $m+1$ rows and $m+1$ columns of Q_m. B is a $(m+1) \times 1$ zero column vector. $C = \left(\frac{\gamma_{VH}}{a_V+\gamma_{VH}}, 0, 0, \cdots, \frac{a_V}{a_V+\gamma_{VH}}\right)$ with size $1 \times (m+1)$. D is a 1×1 zero matrix. From the construction of Q_m, we can see that $A = Q_{m-1}$. Then

$$Q_m^2 = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A^2 & 0 \\ CA & 0 \end{pmatrix}$$

$$
\begin{aligned}
Q_m^3 &= \begin{pmatrix} A^2 & 0 \\ CA & 0 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A^3 & 0 \\ CA^2 & 0 \end{pmatrix} \\
&\cdots \\
Q_m^m &= \begin{pmatrix} A^{m-1} & 0 \\ CA^{m-1} & 0 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A^m & 0 \\ CA^{m-1} & 0 \end{pmatrix}
\end{aligned} \tag{2.2.64}
$$

We have $A = Q_{m-1}$, so $A^m = Q_{m-1}^m = Q_{m-1}^{m-1} = A^{m-1}$, which is known by induction. Thus

$$
\begin{aligned}
CA^{m-1} &= \left(\frac{\gamma_{VH}}{a_V + \gamma_{VH}}, 0, 0, \cdots, \frac{a_V}{a_V + \gamma_{VH}} \right) Q_{m-1}^{m-1} \\
&= \left(1 - \left(\frac{a_V}{a_V + \gamma_{VH}} \right)^m, \left(\frac{a_V}{a_V + \gamma_{VH}} \right)^m, 0, \cdots, 0 \right)
\end{aligned} \tag{2.2.65}
$$

Substitute equation (2.2.65) into equation (2.2.64), we get Q_m^m equals what we have stated in (2.2.61). Hence the system has a stationary state with the probability $P_H = \left(\frac{a_V}{a_V+\gamma_{VH}}\right)^m$ of virus extinction and the probability $P_D = 1 - \left(\frac{a_V}{a_V+\gamma_{VH}}\right)^m$ of cell extinction. □

Method 2: Direct solution of master equation for the reduced system

We have also found the analytical solution of the master equation for the reduced system. As before, we start with examples in which the initial number of viruses is 1, 2, and 3.

Case 1: $m = 1$ The system (2.2.47) becomes:

$$
\begin{aligned}
\frac{dP_2}{dt} &= \gamma_{VH} P_1 \\
\frac{dP_0}{dt} &= a_V P_1 \\
\frac{dP_1}{dt} &= -(a_V + \gamma_{VH}) P_1
\end{aligned} \tag{2.2.66}
$$

Let $P = (P_2, P_0, P_1)^T$, then $P(0) = (0, 0, 1)^T$. The eigenvalues of the coefficient matrix on the right hand side of (2.2.66) are $\lambda_1 = 0$, $\lambda_2 = 0$ and $\lambda_3 = -(\gamma_{VH} + a_V)$. The associate eigenvectors are $v_1 = (1, 0, 0)^T$, $v_2 = (0, 1, 0)^T$, and $v_3 = \left(\frac{-\gamma_{VH}}{a_v+\gamma_{VH}}, \frac{-a_V}{a_v+\gamma_{VH}}, 1\right)^T$. So the solution of (2.2.66) can be written as

$$
P(t) = c_1 v_1 + c_2 v_2 + c_3 e^{-(a_V+\gamma_{VH})t} v_3 \tag{2.2.67}
$$

Apply the initial condition to solve c_1, c_2 and c_3. We get

$$c_1 = \frac{\gamma_{VH}}{a_V + \gamma_{VH}}, \quad c_2 = \frac{a_V}{a_V + \gamma_{VH}}, \quad c_3 = 1 \tag{2.2.68}$$

So the solution of (2.2.66) is

$$\begin{aligned} P_2(t) &= \frac{\gamma_{VH}}{a_V + \gamma_{VH}}\left(1 - e^{-(a_V+\gamma_{VH})t}\right) \\ P_0(t) &= \frac{a_V}{a_V + \gamma_{VH}}\left(1 - e^{-(a_V+\gamma_{VH})t}\right) \\ P_1(t) &= e^{-(a_V+\gamma_{VH})t} \end{aligned} \tag{2.2.69}$$

As time t approaches infinity, we have

$$P_2^{limit} = 1 - \frac{a_V}{a_V + \gamma_{VH}}, \quad P_0^{limit} = \frac{a_V}{a_V + \gamma_{VH}}, \quad P_3^{limit} = 0 \tag{2.2.70}$$

The above results agree with the results obtained using EMC. Since we have analytically solved the system, we can calculate the conditional mean extinction time of virus. Let ΔP_0 be the probability change of P_0 in a small time interval, thus $\Delta P_0 = (a_V P_1)\Delta t$. P_0^{limit}, the normalized factor, is the limiting probability that the system ends at healthy state, in the other word, virus go extinction. So the conditional mean virus extinction time can be calculated in the following way:

$$\begin{aligned} \langle t_V \rangle = \frac{1}{P_0^{limit}} \sum t(\Delta P_0) &= \left(\frac{a_V}{a_V + \gamma_{VH}}\right)^{-1} \int_0^\infty t(a_V P_1) dt \\ &= (a_V + \gamma_{VH}) \int_0^\infty t e^{-(a_V+\gamma_{VH})t} dt \\ &= \frac{1}{a_V + \gamma_{VH}} \end{aligned} \tag{2.2.71}$$

The variance of conditional virus extinction time is

$$var_V = \langle t_V^2 \rangle - \langle t_V \rangle^2 = \frac{1}{(a_V + \gamma_{VH})^2} \tag{2.2.72}$$

Case 2: $m = 2$ The system (2.2.47) becomes:

$$\begin{aligned} \frac{dP_3}{dt} &= \gamma_{VH} P_1 + 2\gamma_{VH} P_2 \\ \frac{dP_0}{dt} &= a_V P_1 \end{aligned} \tag{2.2.73}$$

$$\begin{aligned}\frac{dP_1}{dt} &= 2a_V P_2 - (a_V + \gamma_{VH})P_1 \\ \frac{dP_2}{dt} &= -2(a_V + \gamma_{VH})P_2\end{aligned}$$

Let $P = (P_3, P_0, P_1, P_2)^T$, then $P(0) = (0, 0, 0, 1)^T$. The eigenvalues of the coefficient matrix on the right hand side of (2.2.73) are $\lambda_1 = 0$, $\lambda_2 = 0$ and $\lambda_3 = -(a_V + \gamma_{VH})$, $\lambda_4 = -2(a_V + \gamma_{VH})$. The associate eigenvectors are $v_1 = (1, 0, 0, 0)^T$, $v_2 = (0, 1, 0, 0)^T$, and $v_3 = \left(\frac{-\gamma_{VH}}{a_V+\gamma_{VH}}, \frac{-a_V}{a_v+\gamma_{VH}}, 1\right)^T$, $v_4 = \left(\frac{-\gamma_{VH}^2}{(a_V+\gamma_{VH})^2}, \frac{-a_V^2}{(a_v+\gamma_{VH})^2}, \frac{-2a_V}{a_V+\gamma_{VH}}\right)^T$. Similarly as we did in (2.2.66), by applying the initial condition, (2.2.73) can be solved as

$$\begin{aligned}P_3(t) &= c_1 + c_3\frac{-\gamma_{VH}}{a_V + \gamma_{VH}}e^{-(a_V+\gamma_{VH})t} + c_4\frac{-{\gamma_{VH}}^2}{(a_V + \gamma_{VH})^2}e^{-2(a_V+\gamma_{VH})t} \\ P_0(t) &= c_2 + c_3\frac{-a_V}{a_V + \gamma_{VH}}e^{-(a_V+\gamma_{VH})t} + c_4\frac{-{a_V}^2}{(a_V + \gamma_{VH})^2}e^{-2(a_V+\gamma_{VH})t} \\ P_1(t) &= c_3 e^{-(a_V+\gamma_{VH})t} + c_4\frac{-2a_V}{a_V + \gamma_{VH}}e^{-2(a_V+\gamma_{VH})t} \\ P_2(t) &= c_4 e^{-2(a_V+\gamma_{VH})t}\end{aligned} \tag{2.2.74}$$

where the values of the coefficients are

$$\begin{aligned}c_1 &= 1 - \left(\frac{a_V}{a_V + \gamma_{VH}}\right)^2, \quad c_2 = \left(\frac{a_V}{a_V + \gamma_{VH}}\right)^2, \\ c_3 &= \frac{2a_V}{a_V + \gamma_{VH}}, \quad c_4 = 1\end{aligned} \tag{2.2.75}$$

As time t approaches infinity, we have

$$\begin{aligned}P_3^{limit} &= c_1 = 1 - \left(\frac{a_V}{a_V + \gamma_{VH}}\right)^2 \\ P_0^{limit} &= c_2 = \left(\frac{a_V}{a_V + \gamma_{VH}}\right)^2 \\ P_1^{limit} &= P_2^{limit} = 0\end{aligned} \tag{2.2.76}$$

Then the conditional virus extinction time is

$$\begin{aligned}\langle t_V \rangle &= \sum t(\Delta P_0)/P_0^{limit} \\ &= \left(\frac{a_V}{a_V + \gamma_{VH}}\right)^{-2}\int_0^\infty t(a_V P_1)dt\end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{a_V+\gamma_{VH}}{a_V}\right)^2 \int_0^\infty t\left(\frac{2a_V}{a_V+\gamma_{VH}}e^{-(a_V+\gamma_{VH})t} - \frac{2a_V}{a_V+\gamma_{VH}}e^{-2(a_V+\gamma_{VH})t}\right)dt \\
&= \frac{3}{2(a_V+\gamma_{VH})} \qquad (2.2.77)
\end{aligned}
$$

The variance of conditional virus extinction time is

$$
var_V = \langle t_V^2\rangle - \langle t_V\rangle^2 = \frac{5}{4(a_V+\gamma_{VH})^2} \qquad (2.2.78)
$$

Case 3: $m = 3$ The system (2.2.47) becomes:

$$
\begin{aligned}
\frac{dP_4}{dt} &= \gamma_{VH}P_1 + 2\gamma_{VH}P_2 + 3\gamma_{VH}P_3 \\
\frac{dP_0}{dt} &= a_V P_1 \\
\frac{dP_1}{dt} &= 2a_V P_2 - (a_V+\gamma_{VH})P_1 \\
\frac{dP_2}{dt} &= 3a_V P_3 - 2(a_V+\gamma_{VH})P_2 \qquad (2.2.79) \\
\frac{dP_3}{dt} &= -3(a_V+\gamma_{VH})P_3
\end{aligned}
$$

The initial condition is $P(0) = (0,0,0,1)^T$. Similarly, we solve this system by finding its eigenvalues and associated eigenvectors of the coefficient matrix. The solution is below:

$$
\begin{aligned}
P_4(t) &= 1 - P_0(t) - P_1(t) - P_2(t) - P_3(t) \\
P_0(t) &= \frac{a_V^3}{(a_V+\gamma_{VH})^3}\left(1-e^{-(a_V+\gamma_{VH})t}\right)^3 \\
P_1(t) &= \frac{3a_V^2 e^{-(a_V+\gamma_{VH})t}}{(a_V+\gamma_{VH})^2}\left(1-e^{-(a_V+\gamma_{VH})t}\right)^2 \qquad (2.2.80) \\
P_2(t) &= \frac{3a_V e^{-2(a_V+\gamma_{VH})t}}{a_V+\gamma_{VH}}\left(1-e^{-(a_V+\gamma_{VH})t}\right) \\
P_3(t) &= e^{-3(a_V+\gamma_{VH})t}
\end{aligned}
$$

As time t approaches infinity, we have

$$
\begin{aligned}
P_4^{limit} &= 1 - \left(\frac{a_V}{a_V+\gamma_{VH}}\right)^3, \quad P_0^{limit} = \left(\frac{a_V}{a_V+\gamma_{VH}}\right)^3 \qquad (2.2.81) \\
P_1^{limit} &= P_2^{limit} = P_3^{limit} = 0
\end{aligned}
$$

Then the conditional virus extinction time is

$$\langle t_V \rangle = \sum t(\Delta P_0)/P_0^{limit} = \left(\frac{a_V}{a_V+\gamma_{VH}}\right)^{-3}\int_0^{\infty} t(a_V P_1)dt = \frac{11}{6(a_V+\gamma_{VH})} \quad (2.2.82)$$

The variance of conditional virus extinction time is

$$var_V = \langle t_V^2 \rangle - \langle t_V \rangle^2 = \frac{49}{36(a_V+\gamma_{VH})^2} \quad (2.2.83)$$

After these examples, we are ready to formulate the general result:

General Case: The initial number of virus is m. Let $P = (P_{m+1}, P_0, P_1, \ldots, P_m)^T$ be the probability of state $m+1$, 0, 1, ..., m.

Theorem 2.2.2. *When the initial condition of the system (2.2.47) is $P(0) = (0, 0, \ldots, 0, 1)^T$ with size $1 \times (m+2)$, the solution of (2.2.47) is:*

$$\begin{aligned} P_k(t) &= \frac{C_m^k a_V^{m-k}}{(a_V+\gamma_{VH})^{m-k}} e^{-k(a_V+\gamma_{VH})t}(1-e^{-(a_V+\gamma_{VH})t})^{m-k}, \quad k = 0, 1, ..., m \\ P_{m+1}t &= 1 - \sum_{k=0}^{m} P_k(t) \end{aligned} \quad (2.2.84)$$

where $C_m^k = \frac{m!}{(m-k)!k!}$. The conditional mean virus extinction time is given by

$$\langle t_V \rangle = \frac{1}{(a_V+\gamma_{VH})} \sum_{i=1}^{m} \frac{1}{i} \quad (2.2.85)$$

and the variance of conditional virus extinction time is given by

$$var_V = \frac{1}{(a_V+\gamma_{VH})^2} \sum_{i=1}^{m} \frac{1}{i^2} \quad (2.2.86)$$

Proof. The result can be proven by direct substitution. When $k = 0$,

$$\begin{aligned}\dot{P}_0(t) &= \left(\frac{a_V^m}{(a_V+\gamma_{VH})^m}(1-e^{-(a_V+\gamma_{VH})t})^m\right)' \\ &= \frac{a_V^m}{(a_V+\gamma_{VH})^{m-1}} m e^{-(a_V+\gamma_{VH})t}(1-e^{(-a_V+\gamma_{VH})t})^{m-1} \\ &= a_V P_1(t) \end{aligned} \tag{2.2.87}$$

When $k = 1, 2, ..., m-1$,

$$\begin{aligned}\dot{P}_k(t) &= \left(\frac{C_m^k a_V^{m-k}}{(a_V+\gamma_{VH})^{m-k}} e^{-k(a_V+\gamma_{VH})t}(1-e^{-(a_V+\gamma_{VH})t})^{m-k}\right)' \\ &= \frac{C_m^k a_V^{m-k}}{(a_V+\gamma_{VH})^{m-k}}(a_V+\gamma_{VH})e^{-k(a_V+\gamma_{VH})t}\left(-k(1-e^{-(a_V+\gamma_{VH})t})^{m-k}\right) \\ &\quad + \frac{C_m^k a_V^{m-k}}{(a_V+\gamma_{VH})^{m-k}}(a_V+\gamma_{VH})e^{-k(a_V+\gamma_{VH})t}(m-k)e^{-(a_V+\gamma_{VH})t}(1-e^{-(a_V+\gamma_{VH})t})^{m-k-1} \\ &= -k(a_V+\gamma_{VH})P_k(t) + \frac{C_m^k a_V^{m-k}(m-k)}{(a_V+\gamma_{VH})^{m-k-1}} e^{-(k+1)(a_V+\gamma_{VH})t}(1-e^{-(a_V+\gamma_{VH})t})^{m-k-1} \\ &= -k(a_V+\gamma_{VH})P_k(t) + (k+1)a_V P_{k+1}(t) \end{aligned} \tag{2.2.88}$$

When $k = m$,

$$\dot{P}_m(t) = \left(e^{-m(a_V+\gamma_{VH})t}\right)' = -m(a_V+\gamma_{VH})P_m(t) \tag{2.2.89}$$

Hence formulas in (2.2.84) give the solution of (2.2.47).

The conditional virus mean extinction time can be deduced as follows:

$$\begin{aligned}\langle t_V\rangle &= \frac{1}{P_0^{limit}}\int_0^\infty t(a_V P_1)dt \\ &= \left(\frac{a_V+\gamma_{VH}}{a_V}\right)^m \int_0^\infty t a_V \frac{m a_V^{m-1} e^{-(a_V+\gamma_{VH})t}}{(a_V+\gamma_{VH})^{m-1}}\left(1-e^{-(a_V+\gamma_{VH})t}\right)^{m-1} dt \\ &= m\int_0^\infty t(a_V+\gamma_{VH})e^{-(a_V+\gamma_{VH})t}\left(1-e^{-(a_V+\gamma_{VH})t}\right)^{m-1} dt \\ &= t\left(1-e^{-(a_V+\gamma_{VH})t}\right)^m\Big|_{t=0}^\infty - \int_0^\infty \left(1-e^{-(a_V+\gamma_{VH})t}\right)^m dt \\ &= \int_0^\infty dt - \int_0^\infty \left(1-e^{-(a_V+\gamma_{VH})t}\right)^m dt \\ &= \int_0^\infty dt - \int_0^\infty \left(\left(1-e^{-(a_V+\gamma_{VH})t}\right)^{m-1} - \left(1-e^{-(a_V+\gamma_{VH})t}\right)^{m-1} e^{-(a_V+\gamma_{VH})t}\right) dt \\ &= \int_0^\infty dt + \frac{1}{m(a_V+\gamma_{VH})} - \int_0^\infty \left(1-e^{-(a_V+\gamma_{VH})t}\right)^{m-1} dt \end{aligned}$$

$$
\begin{aligned}
&\vdots \\
&= \frac{1}{(a_V+\gamma_{VH})}\left(\frac{1}{m}+\frac{1}{m-1}+\ldots+\frac{1}{2}\right)+\int_0^\infty dt-\int_0^\infty\left(1-e^{-(a_V+\gamma_{VH})t}\right)dt \\
&= \frac{1}{(a_V+\gamma_{VH})}\left(\frac{1}{m}+\frac{1}{m-1}+\ldots+\frac{1}{2}\right)+\int_0^\infty e^{-(a_V+\gamma_{VH})t}dt \\
&= \frac{1}{(a_V+\gamma_{VH})}\left(\frac{1}{m}+\frac{1}{m-1}+\ldots+\frac{1}{2}+1\right) \\
&= \frac{1}{(a_V+\gamma_{VH})}\sum_{i=1}^{m}\frac{1}{i} \qquad (2.2.90)
\end{aligned}
$$

The formula for the second moments of conditional virus extinction time is:

$$
\begin{aligned}
\left\langle t_V^2\right\rangle &= \frac{1}{P_0^{limit}}\int_0^\infty t^2(a_V P_1)dt \\
&= \left(\frac{a_V+\gamma_{VH}}{a_V}\right)^m\int_0^\infty t^2 a_V\frac{m a_V^{m-1}e^{-(a_V+\gamma_{VH})t}}{(a_V+\gamma_{VH})^{m-1}}\left(1-e^{-(a_V+\gamma_{VH})t}\right)^{m-1}dt \\
&= m\int_0^\infty t^2(a_V+\gamma_{VH})e^{-(a_V+\gamma_{VH})t}\left(1-e^{-(a_V+\gamma_{VH})t}\right)^{m-1}dt \\
&= t^2\left(1-e^{-(a_V+\gamma_{VH})t}\right)^m\Big|_{t=0}^\infty-\int_0^\infty 2t\left(1-e^{-(a_V+\gamma_{VH})t}\right)^m dt \\
&= \int_0^\infty 2tdt-\int_0^\infty 2t\left(1-e^{-(a_V+\gamma_{VH})}\right)^{m-1}dt+\int_0^\infty 2te^{-(a_V+\gamma_{VH})t}\left(1-e^{-(a_V+\gamma_{VH})t}\right)^{m-1}dt \\
&\vdots \\
&= \int_0^\infty 2tdt-\int_0^\infty 2t\left(1-e^{-(a_V+\gamma_{VH})t}\sum_{i=0}^{m-1}\left(1-e^{-(a_V+\gamma_{VH})t}\right)^i\right)dt \\
&= \sum_{i=0}^{m-1}\int_0^\infty 2te^{-(a_V+\gamma_{VH})t}\left(1-e^{-(a_V+\gamma_{VH})t}\right)^i dt \\
&= \sum_{i=1}^{m}\frac{2}{i(a_V+\gamma_{VH})}\left(\int_0^\infty dt-\int_0^\infty\left(1-e^{-(a_V+\gamma_{VH})t}\right)^i dt\right) \quad \textit{by (2.2.85)} \\
&= \frac{2}{(a_V+\gamma_{VH})^2}\sum_{i=1}^{m}\frac{1}{i}\sum_{k=1}^{i}\frac{1}{k} \qquad (2.2.91)
\end{aligned}
$$

Thus the variance of the conditional virus extinction time is:

$$
\begin{aligned}
var_V &= \left\langle t_V^2\right\rangle-\left\langle t_V\right\rangle^2 \\
&= \frac{2}{(a_V+\gamma_{VH})^2}\sum_{i=1}^{m}\frac{1}{i}\sum_{k=1}^{i}\frac{1}{k}-\left(\frac{1}{(a_V+\gamma_{VH})}\sum_{i=1}^{m}\frac{1}{i}\right)^2 \\
&= \frac{1}{(a_V+\gamma_{VH})^2}\sum_{i=1}^{m}\frac{1}{i^2} \qquad (2.2.92)
\end{aligned}
$$

□

Recall that the Gillespie's direct method computes the probabilities and conditional mean time of extinction numerically when n is large. If we take $V_0 = 3$, the trajectories using the Gillespie algorithm reveal that the limits of virus and cell extinction probabilities are 0.36 and 0.64 and the limit of conditional mean virus extinction time is 0.76, see Figure 8. Let's substitute parameter values into the equation (2.2.84) and let $t \to \infty$. We obtain the mean extinction probabilities of virus and cell:

$$P_0^{limit} = \left(\frac{1.7}{1.7+0.7}\right)^3 = 0.36, \quad P_4^{limit} = 1 - P_0^{limit} = 0.64 \tag{2.2.93}$$

We can also substitute the parameters in equation (2.2.85) of mean virus extinction time:

$$\langle t_V \rangle = \frac{1}{(1.7+0.7)}\left(\frac{1}{3} + \frac{1}{2} + 1\right) = 0.76 \tag{2.2.94}$$

We can see that all these values agree with the numerical results. In Figure 10 shows the variance when $V_0 = 1$. By applying formula (2.2.86), we have $var_V = 1/2.4^2 = 0.17$, the same as the variance in Gilliespie's, while the variance in Laplace's is a little smaller. This verifies that Gillespie's algorithm is more accurate than Laplace's method when n is large since (2.2.86) is based on the assumption that n is large.

Formula (2.2.85) and (2.2.86) of the mean and variance of virus extinction time in the approximate model are similar to the formulas in the standard birth-death model in Newman's paper [67].

2.2.2.5 Conclusion of the Stochastic Model In the ODE model, Figure 2 and 3 have showed that the system has three different parameter regions. The stable healthy state $(V, H, I) = (0, 1, 0)$ becomes unstable and an additional steady state - chronic state, appears at $n \approx 5$. Then the steady chronic state becomes unstable and the a limit cycle appears through a Hopf bifurcation at $n \approx 23$. The stochastic model also has 3 parameter regions. When $n < 6$, the virus has an extinction probability 1. That probability decreases monotonically when $6 \leq n < 25$, and then it reaches a limiting value at $n \approx 25$.

We have explored three methods of analysis of the stochastic model: by simulation of the trajectories by the algorithm of Gillespie's direct method, by solving mean extinction times through Laplace transform of the master equation, and by generating formulas for limiting

probability of healthy and dead states and the moments of virus extinction time using a reduced Markov chain.

Gillespie's direct method is capable of dealing with large systems as the required computer memory depends linearly on number of species. Its disadvantage is that many trajectories do not converge to absorbing states and long time simulations are needed for an accurate estimation of extinction times. The Laplace method is fast and accurate. But it can be used only when the system size is small. It requires more memory than Gillespie algorithm when dealing with the same system. The size of the truncated master equation suffers from an exponential growth for an increasing number of chemical species. Elf [26] states that when number of reactions is large, there is no other alternative than a Monte Carlo method.

Other method in the field of stochastic modeling utilizes approximation of the master equations by a continuous Fokker-Planck equation (FPE), which is a partial differential equation in time and N-dimensional space. FPE is well suited for problems in low dimensions especially if high accuracy is desirable. But it is only applicable to problem with few molecular species. Sjoberg [77] states that FPE is less accurate than Gillespie's direct method when the species number is greater than 2. Doering [21] found FPE approximation is valid only near the threshold. Other approximation methods, i.e. chemical Langevin equation, are discussed by Gillespie [33].

2.3 HUMAN IMMUNE RESPONSE

The analysis of the basic system in the previous section shows that in the absence of any immune response, an infected organism is able to recover only for extremely inefficient viruses. The two other possible outcomes are rapid convergent to a dead state and a persistent infection for weak viruses that will approach the dead state eventually. In this section, we show that the immune response makes the disease manageable by:

- isolating the system trajectories from the states that lead to the absorbtion in the dead state
- lowering the total damage to the system,
- shortening the duration of infection (i.e., shortening the mean time of virus extinction),
- increasing the virus extinction probability.

As we mentioned earlier, the immune response to influenza virus has three components: innate, cellular, and adaptive [3, 42]. The innate immune component is localized in the place of infection and is primarily mediated by macrophages and infected cells producing interferons which make susceptible cells resistant to infection. In cellular immune component, signals about the presence of the virus are carried to lymph nodes by antigen presenting cells. Effectors cells pick up on that signal, mature, and migrate to the site of infection, where they neutralize infected cells. In the adaptive immune component, antigens are carried to bone marrow where they are recognized by B-cells. These mature into plasma cells which produce antibody specific to the antigen. The antibodies help to neutralize virus and offer protection from future infection by this same virus. We shall explore the role of each negative feedback separately and in combination with the others mathematically. Figure 16 shows mean extinction times and probabilities in the basic system by the algorithm of Gillespie's direct method. Because the system's behavior changes qualitatively between $n = 6$ and 8, we call $n^* = 6$ a bifurcation of the system. For $n \leq n^*$ the virus extinction probability is 1 and the cell extinction probability is 0, i.e., virus extinction is assured.

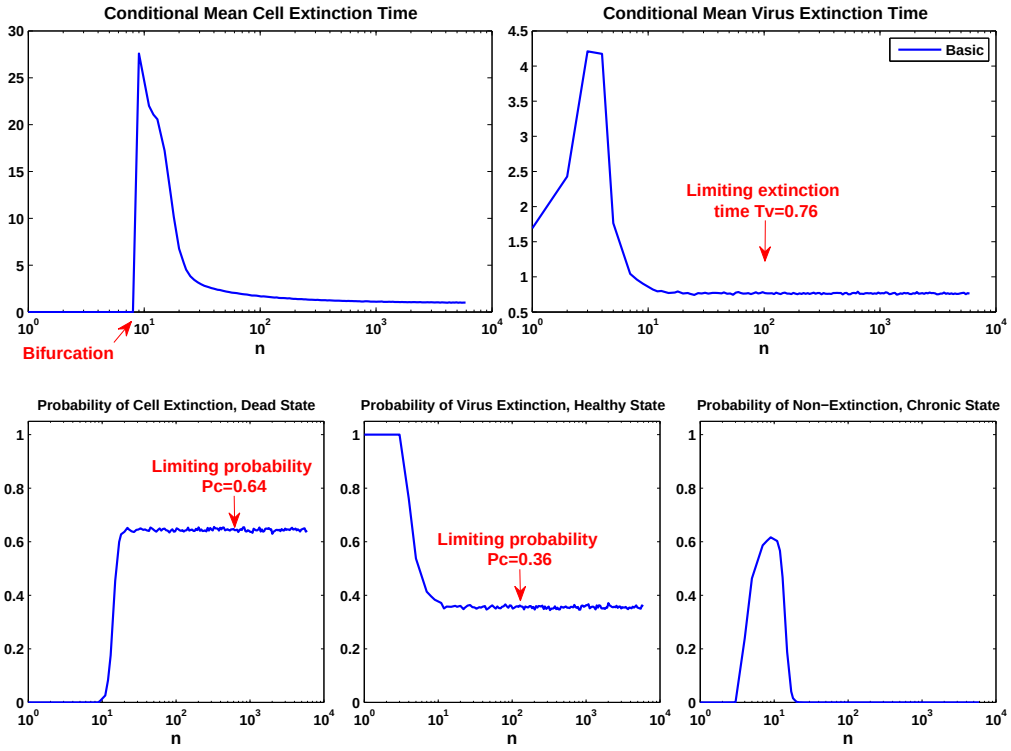

Figure 16: Conditional mean times and probabilities of extinctions in the basic system, $H^* = 100$, $V_0 = 3$. All the trajectories are averaged over 50,000 independent samples generated by Gillespie algorithm. The x-axis is $n = \gamma_V / a_I$ and it is in log scale.

2.3.1 Systems with an Innate Immune Component

The term innate immune component describes a nonspecific response of the organism to a viral infection that is generally mediated by α and β interferon molecules produced by antigen presenting cells which detect the presence of the virus in the system. Interferons interact with the healthy cells and convert them into a resistant state, thereby preventing the virus from spreading efficiently. Mathematically, the innate immune component provides a negative feedback that lowers the virus level. To model the innate response discussed in Hancioglu's paper [40], we assumed the presence of three additional components to the system: the activated antigen presenting cells (**M**) , the interferon (**I**), and the resistant cells (**R**). (Here

we modify the model in Hancioglu's paper [40] by assuming that the interferon molecules are generated solely by infected cells, without any contribution from antigen presenting cells.) The scheme of the corresponding modified system is:

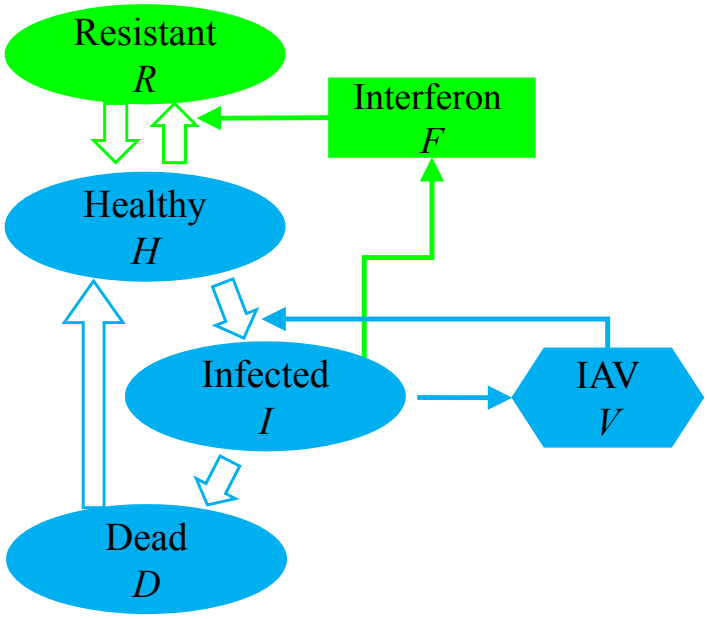

Figure 17: Systems with an innate immune component

The reaction network of the innate immune component is:

$$
\begin{aligned}
V + H &\xrightarrow{\gamma_{VH}} I \\
I &\xrightarrow{a_I} nV + D \\
V &\xrightarrow{a_V} \varnothing \\
H + D &\xrightarrow{b_{HD}} 2H \\
I &\xrightarrow{b_F} I + F && (2.3.1)\\
H + F &\xrightarrow{b_{FH}} R && (2.3.2)\\
F &\xrightarrow{a_F} \varnothing && (2.3.3)\\
R &\xrightarrow{a_R} H && (2.3.4)
\end{aligned}
$$

The presence of resistant cells gives rise to new terms in the equation for healthy cells corresponding to the reproduction of resistant cells, loss of resistance, and conversion of

healthy cells into resistant ones. Antigen presenting cells are simulated by the damage (i.e., dead remnants of infected cells, $D = 1 - I - R - H$), and produce interferon. Interferon molecules interact with healthy cells and convert them into a resistant state. Resistant cells naturally loose resistance and also can reproduce to give rise to healthy cells.

The analysis of influenza A infection in humans [40] yielded the following estimates for the additional parameters. F is re-scaled by 1700 and some other parameters are adjusted by the homeostatic level in our Gillespie algorithm: $b_F = 125$, $b_{FH} = 17/H^*$, $a_F = 8$, $a_R = 1$. The healthy absorbing state of the system is $(V, H, I, R, F) = (0, H^*, 0, 0, 0)$ and dead absorbing state is $(V, H, I, R, F) = (0, 0, 0, 0, 0)$.

Recall in the basic ODE system, there is a limit cycle as γ_V increases. The presence of the innate immune component preserves the structure of the basic diagram (i.e. healthy state stable for small γ_V, then branching into a chronic state) with the exception of missing limit cycle. This implies that innate immune component stabilizes the chronic state for large values of γ_V. (Marissa Goldrich and David Swigon, unpublished result.)

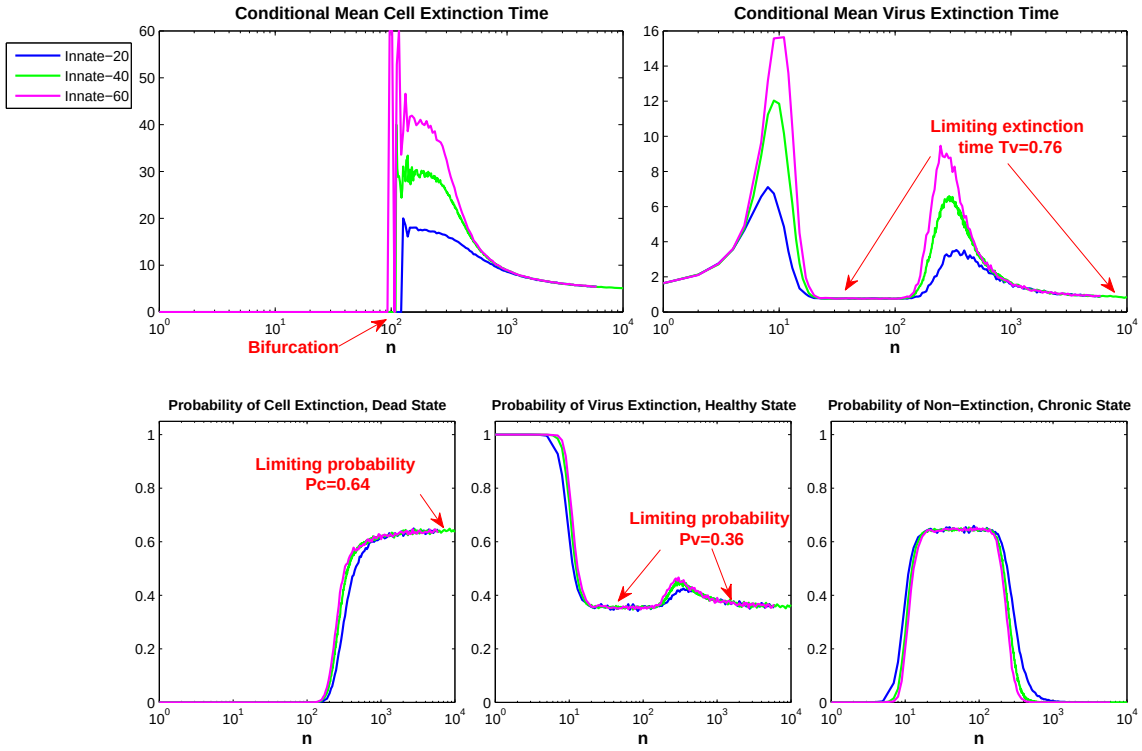

Figure 18: Conditional mean times and probabilities of extinctions in the system with innate immune component. $H^* = 100$, $V_0 = 3$ and all the trajectories are averaged over 50,000 independent samples generated by Gillespie algorithm. The x-axis is $n = \gamma_V/a_I$. The blue, green, and pink trajectories represent the maximum simulation time $T_{max} = 20$, 40 and 60, correspondingly.

Figure 18 is generated by the Gillespie algorithm with $H^* = 100$ and $V_0 = 3$ (Laplace transform method is not applicable to the analysis of the system due to its size.) Compared with the basic system, the innate immune component results in a large increase in n^*. The bifurcation occurs at $n^* = 120$ with the innate immune component included. In basic system, the virus extinction probability reaches the limiting probability when $n > 25$ while with the innate immune component, it doesn't reach the limit until $n > 10^3$. In the plots of conditional mean virus extinction time and probability, we notice that both curves rebound when $n > 200$. This is because what we calculate is the *conditional* mean extinction time.

When $10 < n < 200$, a large part of the simulated trajectories are of type III (no-extinction category), corresponding to a chronic infection. Those type III trajectories lead to an underestimation of the conditional mean extinction time. If we run the simulation long enough, we expect these type III trajectories to end in either virus extinction or cell extinction. We have run the trajectories with the maximum simulation time $T_{max} = 20$, 40, and 60. Figure 18 shows that the longer maximum simulation time results a higher peak of the conditional mean extinction time. Ideally, the curve of virus extinction probability should decrease monotonically until it reaches the limiting probability. When n is between 20 to 120, the two flat lines in virus extinction time and probability are equal to the limiting extinction time $t_V = 0.76$ and probability $P_V = 0.36$, respectively. The reduced Markov chain in limiting case of innate component is the same as that of the basic system. The limiting extinction probabilities here are also: 0.64 for cell extinction and 0.36 for the virus extinction.

2.3.2 Systems with a Cellular Immune Component

The cellular immune component refers to the mechanism of a specific response in which infected cells are neutralized by the action of effector cells (usually CTL's or natural killer cells). This mechanism reduces the source of new virus particles and hence constitutes another form of a negative feedback. The additional components of the system are antigen presenting cells (**M**) and effector cells (**E**). The scheme of this system is:

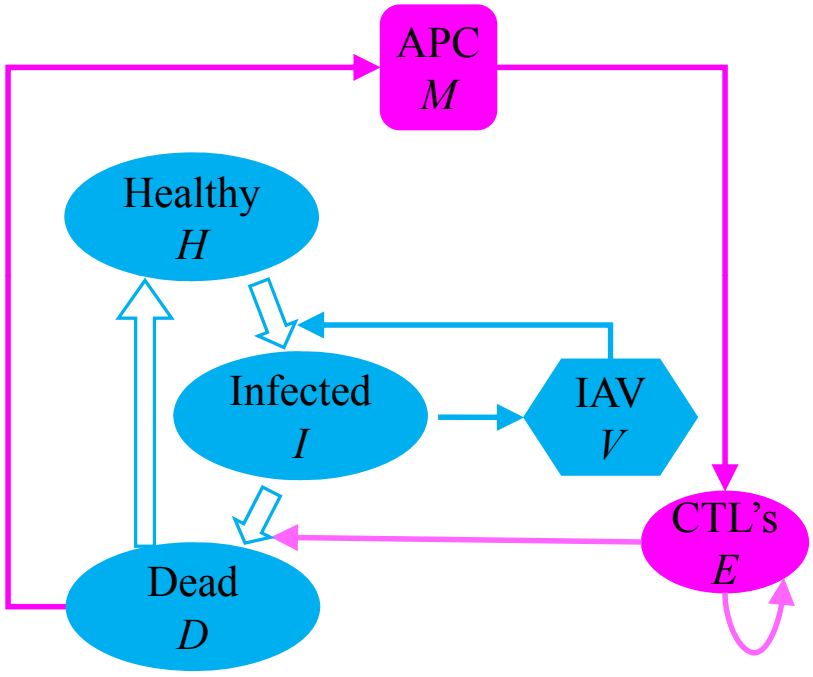

Figure 19: Systems with a cellular immune component

The reaction network of the cellular immune component is:

$$
\begin{aligned}
V + H &\xrightarrow{\gamma_{VH}} I \\
I &\xrightarrow{a_I} nV + D \\
V &\xrightarrow{a_V} \varnothing \\
H + D &\xrightarrow{b_{HD}} 2H \\
D + N &\xrightarrow{b_{MD}} D + M && (2.3.5)\\
M &\xrightarrow{a_M} N && (2.3.6)\\
M &\xrightarrow{b_{EM}} M + E && (2.3.7)\\
E + I &\xrightarrow{b_{EI}} D && (2.3.8)\\
\varnothing &\xrightarrow{b_E} E && (2.3.9)\\
E &\xrightarrow{a_E} \varnothing && (2.3.10)
\end{aligned}
$$

Antigen presenting cells are stimulated by the damage (i.e., dead remnants of infected cells, here $D = 1 - I - H$), the production of effector cells is controlled by the APCs and the effector cells interact with the infected cells and neutralize them, with $p = (b_{IE}H^*)/(b_{EI}E^*)$ the average amount of infected cells destroyed by one effector cell. Reaction (2.3.9) and (2.3.10) correspond to a homeostatic maintenance effect by which the organism strives to maintain the population of effector cells at a homeostatic level. Effector cell E is re-scaled as the multiple of the homeostatic level $E^* = 6 \times 10^7$. The analysis of influenza A infection in Hancioglu's paper [40] yielded the following estimates for the additional parameters (some of the parameters are adjusted by the homeostatic level in our Gillespie algorithm): $b_{MD} = 2/H^*$, $a_M = 0.5$, $b_{EM} = 8.3E^*/M^*$, $b_{EI} = 4.72/H^*$, $b_E = 0.4H^*$, 0.4. The healthy state of the system is $(V, H, I, N, M, E) = (0, H^*, 0, M^*, 0, E^*)$ and dead absorbing states are $(V, H, I, N, M, E) = (0, 0, 0, ND, MD, ED)$.

Figure 20 is generated by the Gillespie algorithm with $H^* = 100$, $M^* = 20$, $E^* = 10$ and $V_0 = 3$. The system with cellular immune component has the bifurcation at $n^* = 10$ while the basic system has $n^* = 6$. The virus extinction time reaches the limit when $n = 50$ in this system compared with $n = 12$ in basic system. It seems cellular immune component doesn't have much improvement at these two n values compared with the innate immune component. Its most important advantage is that the limiting probability of virus extinction is increased from 0.36 to 0.47, and correspondingly, the limiting probability of cell extinction is decreased from 0.64 to 0.53.

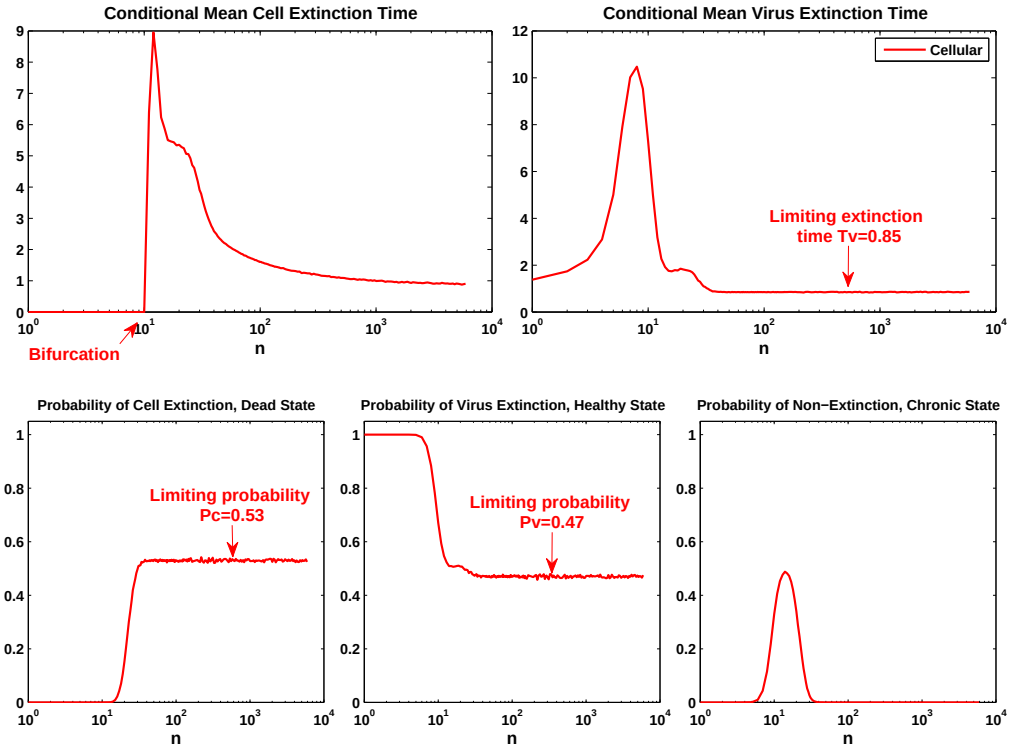

Figure 20: Conditional mean times and probabilities of extinctions in the system with cellular immune component. $H^* = 100$, $M^* = 2$, $E^* = 10$, $V_0 = 3$, and all the trajectories are averaged over 50,000 independent samples generated by Gillespie algorithm. The x-axis is $n = \gamma_V / a_I$.

We have constructed an approximate Markov Chain to derive the limiting probability of this system - with cellular immune component.

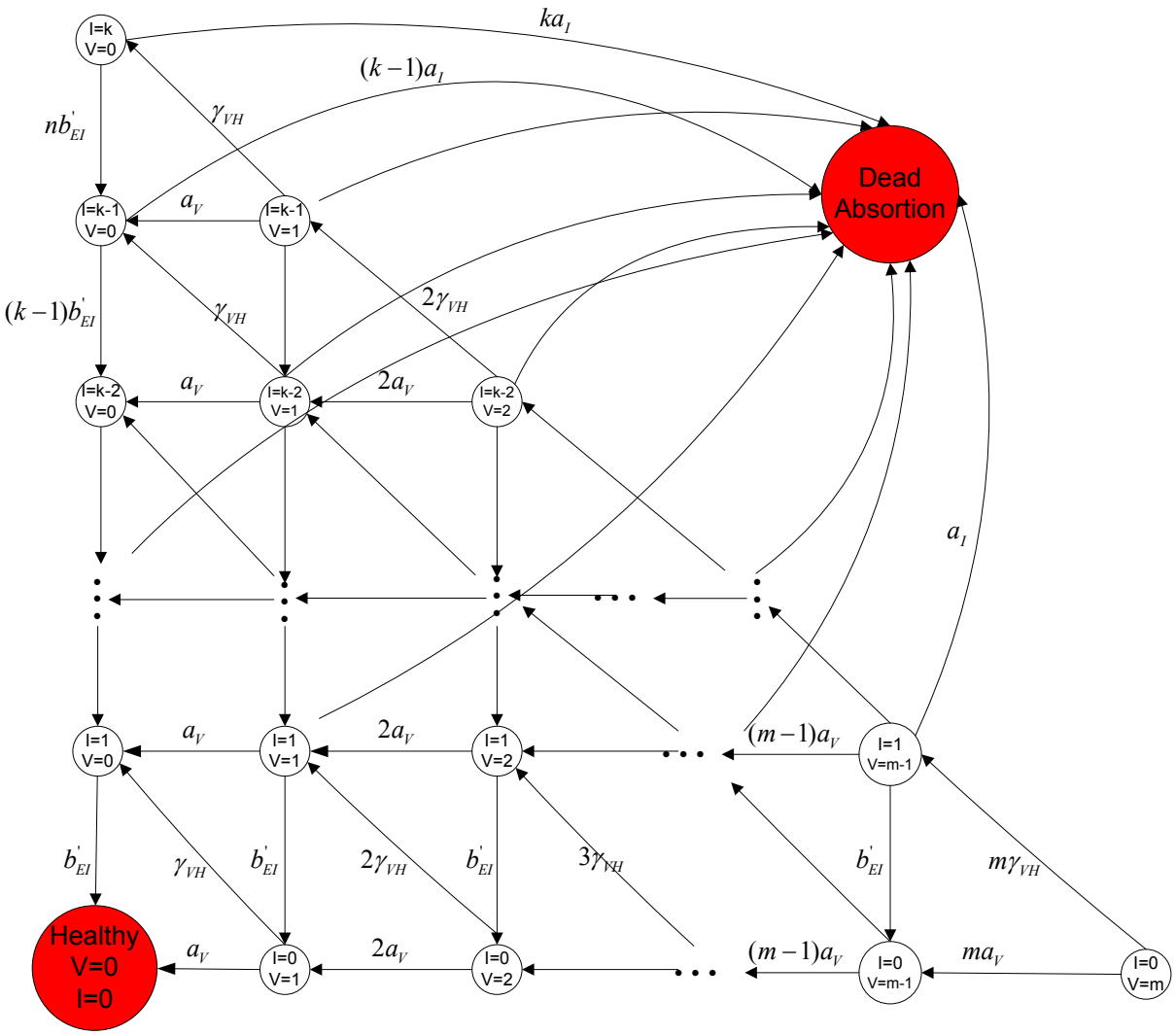

Figure 21: Schema of the approximate Markov chain in the system with a cellular immune component, $b'_{EI} = b_{EI}/E^*$.

Figure 21 is the simplified Markov chain of the system with a cellular immune component. Recall in the approximate model of the basic system, if n is large, once the reaction $V + H \xrightarrow{\gamma_{VH}} I$ occurs, then the system enters the dead absorbing region. We shall use the same approximation here. The difference is that in the system with cellular immune component, the infected cell can be cleared by $E+I \xrightarrow{b_{EI}} D$. So there is two ways to approach the healthy state - from state $(V, I) = (1, 0)$ and state $(V, I) = (0, 1)$. Enumerate all the states in the

Figure 21 and find the limit of the transition matrix. The limiting probability of absorbing states are solved through the embedded method [78], which we have discussed in the basic system.

Theorem 2.3.1. *In the system with a cellular immune component, when $n \to \infty$, the limiting probability P_H of the healthy state and the limiting probability P_D of the dead state are*

$$P_H = \left(\frac{a_I a_V + b'_{EI}(a_V + \gamma_{VH})}{a_I(a_V + \gamma_{VH}) + b'_{EI}(a_V + \gamma_{VH})} \right)^{V_0} \tag{2.3.11}$$

$$P_D = 1 - P_H \tag{2.3.12}$$

where $b'_{EI} = b_{EI}/E^*$.

Proof. This can be proved using the embedded Marko Chain method [78], similarly as we did in Theorem 2.2.1. □

2.3.3 Systems with Combined Innate and Cellular Immune Components

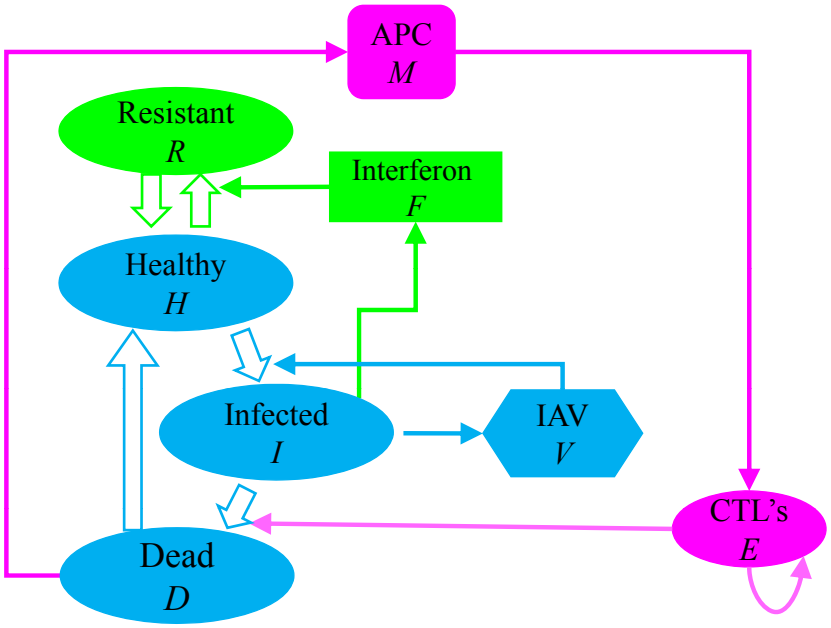

Figure 22: Systems with combined innate and cellular immune Components

The combination of innate and cellular immune components leads to postponement of the bifurcation to larger value of n and an increase in the virus extinction probability (see Figure 23). The bifurcation occurs at $n^* = 200$. The rebound in the virus extinction probability remains because of the limited simulation time. If we run the simulation long enough, the chronic states will terminate primarily in the dead state. The curve of the virus extinction probability should decrease smoothly to the limiting probability. The system has a healthy state at $(V, H, I, R, F, N, M, E) = (0, H^*, 0, 0, 0, M^*, 0, E^*)$ and dead absorbing states that satisfy $(V, H, I, R, F, N, M, E) = (0, 0, 0, 0, 0, ND, MD, ED)$. The reduced Markov chain in limiting case here is the same as the one in the system with only cellular immune response with limiting probabilities of cell extinction $P_C = 0.53$ and virus extinction $P_V = 0.47$.

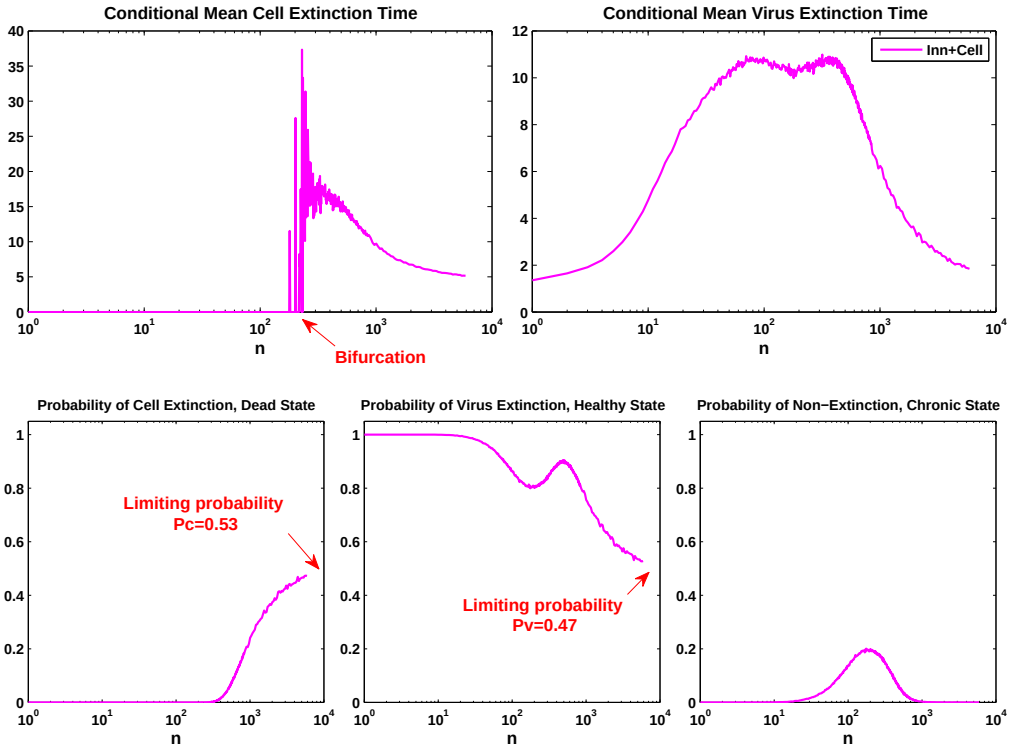

Figure 23: Conditional mean times and probabilities of extinction in system with combined innate and cellular immune components. $H^* = 100$, $M^* = 20$,$E^* = 10$,$V_0 = 3$ and all the trajectories are averaged over 50,000 independent samples generated by Gillespie algorithm. The x-axis is $n = \gamma_V / a_I$.

2.3.4 Systems with a Complete Immune Response

In this section, we consider the combination of all the three immune components. The effect of adaptive immune component on the stability of the healthy state in the ODE model was discussed at length in Hancioglu's paper [40]. For completeness, adaptive immune component helps in clearing the infection by producing antibodies that neutralize the virus particles by rendering them unable to infect healthy cells. The antibodies are produced by plasma cells, production of which is stimulated by APCs. The new components of the system are plasma cells (**P**) and antibodies (**A**).

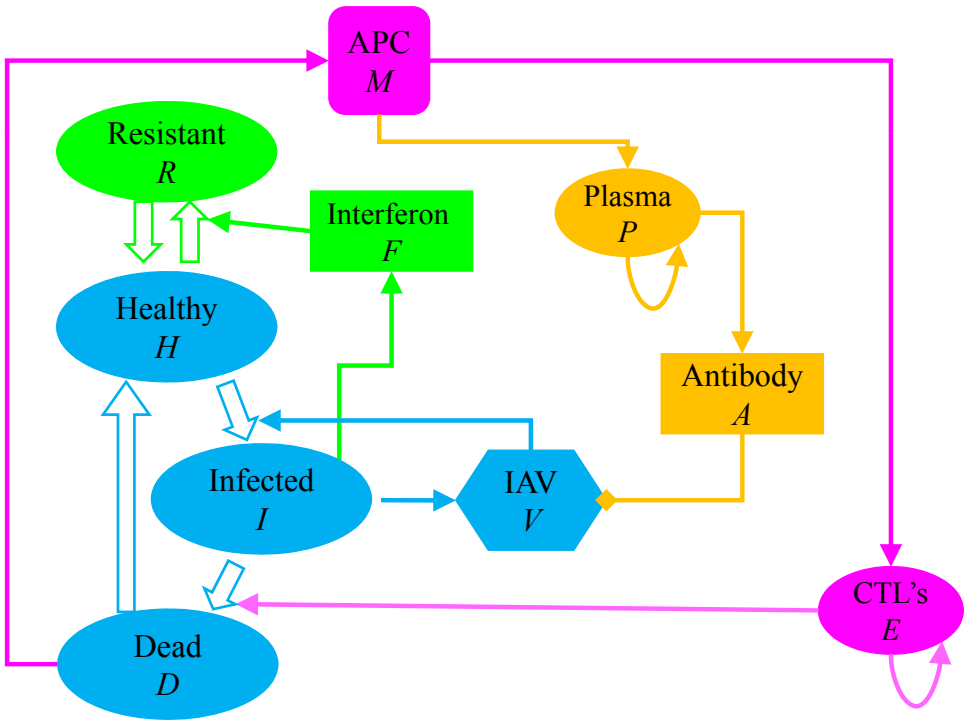

Figure 24: Human immunity system - combined innate, cellular and adaptive immune components

The new reactions added here are:

$$M \xrightarrow{b_{PM}} M + P \tag{2.3.13}$$

$$\varnothing \xrightarrow{a'_P} P \tag{2.3.14}$$

$$P \xrightarrow{a_P} \varnothing \tag{2.3.15}$$

$$P \xrightarrow{b_{PA}} P + A \tag{2.3.16}$$

$$V + A \xrightarrow{\gamma_{AV}} \varnothing \tag{2.3.17}$$

$$A \xrightarrow{a_A} \varnothing \tag{2.3.18}$$

The production of plasma cells is stimulated by activated macrophages and the production of antibodies is stimulated by plasma cells. Antibodies remove viruses and have natural decays. The parameter corresponded to antigenic compatibility of antigenic compatibility of antibodies is assumed to be constant. The parameter values are again estimated from

Hancioglu's paper [40]. We have adjusted some values by the homeostatic level in our Gillespie algorithm: $b_{PM} = 11.5P^*/M^*$, $a'_P = 0.4P^*$, $a_P = 0.4$, $b_{PA} = 0.043A^*/P^*$, $\gamma_{VA} = 14.6/H^*$, $a_A = 0.043$. The system has a healthy state at $(V, H, I, R, F, N, M, E, P, A) = (0, H^*, 0, 0, 0, M^*, 0, E^*, P^*, A^*)$ and dead absorbing states $(V, H, I, R, F, N, M, E, P, A) = (0, 0, 0, 0, 0, ND, MD, ED, PD, AD)$.

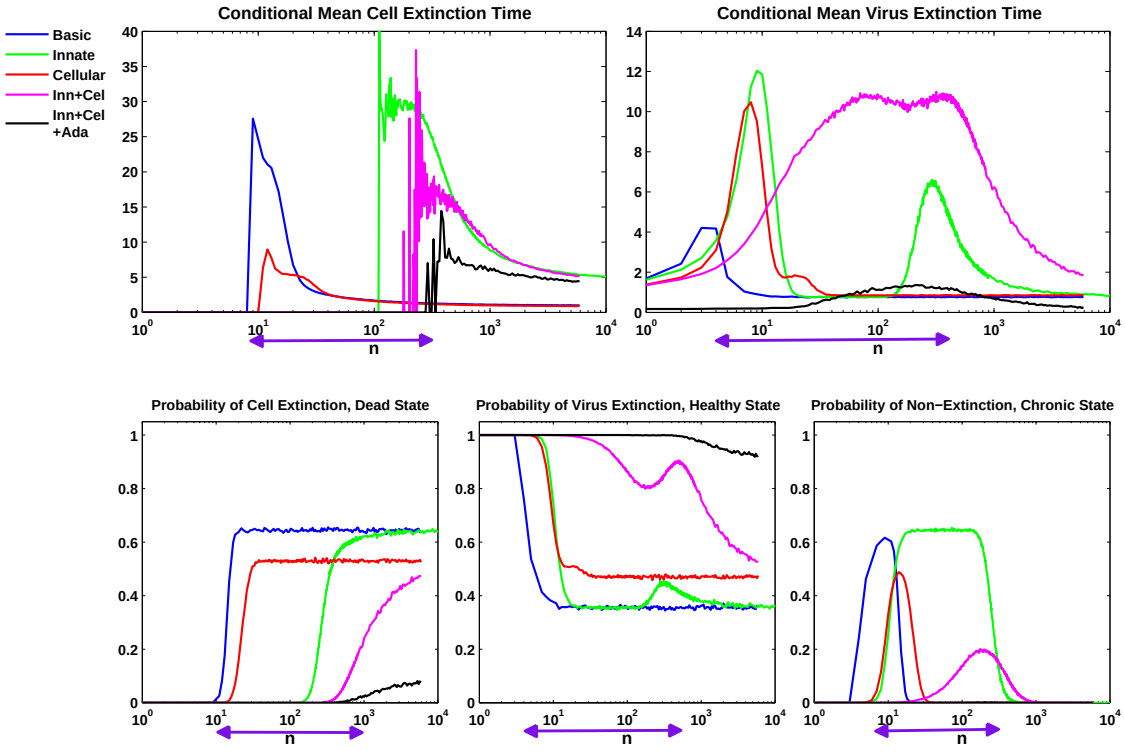

Figure 25: Conditional mean times and probabilities of extinction in human immunity. $H^* = 100$, $M^* = 20$, $E^* = 10$, $P^* = 10$, $A^* = 100$, $V_0 = 3$, and all the trajectories are averaged over 50,000 independent samples generated by Gillespie algorithm. The x-axis is $n = \gamma_V/a_I$. The trajectories of the system with no immune component, with innate immune component, with cellular immune component, with the combination of innate and cellular immune components, and with the complete immune response - the combination of innate, cellular and adaptive immune components are in blue, green, red, pink, and black. The purple arrow represents biologically reasonable values of n.

Figure 25 summarizes the earlier graphs and adds the plots for the combined response. The bifurcation occurs at $n^* \approx 300$ for the black curve. The purple arrow at the bottom of each plot represents biologically reasonable reproduction numbers for viruses. All three types of immune components are essential. In the basic system (blue lines in Figure 25), healthy cells go extinct at small number of n and the organism dies. The innate (green) immune component substantially reduces cell extinction probability when $n < 200$ and promotes chronic disease. The cellular (red) immune component increases the virus extinction probability at any n numbers. Then the combination of two (pink) almost eliminate cell extinction and significantly increase virus extinction rate. The adaptive immune component aids in virus clearance and prevents infection altogether. The conditional mean virus extinction time is shortened substantially. The probability of virus extinction is above 90% across the range of n. The probability of cell extinction is nonzero only for n above 10^3. The combination of the three immune component (black) makes the human body almost invulnerable under virus attack unless a large number of viruses can be generated from a single infected cell.

Theorem 2.3.2. *In the system with innate, cellular and adaptive immune components, when $n \to \infty$, the limiting probability P_H of the healthy state and the limiting probability P_D of the dead state are*

$$P_H = \left(\frac{(a_V + \gamma_{AV} A^*) a_I + (a_V + \gamma_{AV} A^* + \gamma_{VH}) b'_{EI}}{(a_V + \gamma_{AV} A^* + \gamma_{VH})(a_I + b'_{EI})} \right)^{V_0} \tag{2.3.19}$$

$$P_D = 1 - P_H \tag{2.3.20}$$

where $b'_{EI} = b_{EI}/E^*$.

Proof. This can be proved through the embedded method [78], which we have discussed in the basic system. □

3.0 VOCAL FOLD INFLAMMATION

3.1 INTRODUCTION

Human vocal folds are located in the larynx. When air stream from the lungs passes through the vocal folds, they will be set into oscillation and produce a periodic train of air pulses. That is how sound is produced. If the vocal folds are stressed or traumatized by noxious stimuli, the tissue properties of the vocal folds will be altered and the vocal fold oscillation will be disturbed, result in an altered voice quality, i.e., a voice disorder [51].

Phonotrauma

Voice disorders appear to be the most common communication disorders across the human lifespan [74]. In most cases, these disorders are not "cosmetic". They can produce substantial negative effects on quality of life in affected individuals as well as substantial negative economic impact on society [86]. A common cause of voice disorder is phonotrauma, which is a vocal fold injury associated with phonation (i.e., voice production). When people speak or sing, the vocal folds oscillate at a relatively high frequency. Typically, the vocal folds make contact at some point during the vibratory cycle. The associated motions generate various types of bio-mechanical forces within the oscillating vocal fold tissues [37, 38, 43, 80]. Although micro-injury likely occurs within the tissue every day, injury may or may not be functionally relevant, depending on the magnitude of the injury and the host's requirement for an intact voice. Specifically, under conditions of typical voice use, the tissues are generally competent to withstand the bio-mechanical forces physiologically, probably with relatively minor and transient local inflammatory responses. In such cases, the levels of tissue injury are sufficiently microscopic to be regenerated intrinsically without a full scale inflammatory or wound healing response. On the other hand, when phonation

conditions are frankly phonotraumatic, the associated bio-mechanical forces are sufficiently large to cause substantial cell death and macroscopic tissue damage. If such conditions persist without adequate rest, damaged tissue does not have time needed to repair [81]. As a result, benign lesions at the mid-membranous vocal folds may develop, leading to dysphonia and voice-related changes in quality of life [58].

Therapeutic management of phonotrauma may involve medical or sometimes surgical intervention. However, behavioral voice therapy is generally considered central to any treatment program, and is usually preferred as the first-line approach [6, 14, 44]. Unfortunately, strong scientific evidence is still lacking regarding the effectiveness and efficacy of voice therapy for phonotrauma. One of the research obstacles is that the underlying mechanism of inflammation and healing underlying phonotrauma is highly complex. Common research approaches, such as randomized clinical trials, are not sufficient to capture the dynamics of the disease and its response to treatments [51]. In this study, a systems biology approach utilizing computational modeling was used to simulate acute phonotrauma and its response to behavioral treatments.

Development of mathematical models of inflammation and healing

Inflammation and healing, which are innate tissue responses to mechanical stress/ trauma, are regulated by a complex dynamic system. A modeling approach, which combines empirical, mathematical and computational tools, was taken to study the biological complexity of this dynamic system.

Computational agent-based models (ABMs) were developed to quantitatively characterize multiple cellular and molecular interactions around inflammation and healing. Initially, the ABMs of the vocal fold injury were specified to the setting of phonotrauma, a biomechanical injury owing to phonation. These models allowed for tests of various hypothetical effects of motion-based treatments in individuals with acute phonotrauma [53, 51]. The phonotrauma ABMs were calibrated and validated with empirical data of a panel of inflammatory mediators, obtained from laryngeal secretions in individuals following experimentally induced phonotrauma and a randomly assigned motion-based treatment. In addition, an ABM of surgically induced vocal fold trauma was developed and subsequently calibrated to empirical data of inflammatory mediators and extracellular matrix substances from rat stud-

ies [51]. In general, ABM simulations reproduced and predicted trajectories of inflammatory mediators and extracellular matrix as seen in empirical data of phonotrauma and surgical vocal fold trauma. The simulation results illustrated a spectrum of inflammatory responses to phonotrauma, surgical trauma and motion-based treatments.

In addition to the use of ABMs for simulating vocal fold injuries [53, 51], the parallel development of an ordinary differential equation (ODE) model for phonotrauma was pursued in the interest of cross-platform comparison of results [52]. A rudimentary ODE was developed using the module of system dynamics included in the Netlogo platform (Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL). To test the validity of our agent-based model of phonotrauma, "model docking" was used. "Model docking" is a well-vetted validation strategy based on a comparison of predictions from different models across an array of user input data. The finding of similar predictions in agent-based and equation-based models would increase confidence in the underlying assumptions made in the previous ABM. In this comparison study, the ABM and ODE model predicted similar cellular and molecular patterns for inflammatory and wound healing responses under small initial damage. However, the models' results diverged in their predictions of inflammatory and wound healing responses for a large initial insult. The other limitation was that the module of system dynamics in Netlogo was just released at the time of the project. The module was limited for conventional mathematical analysis, such as, parameter fitting and sensitivity tests.

Therefore, in the current study, we adopted the well-established MATLAB platform to rigorously pursue the development of ODE models in simulating inflammation and healing in the setting of vocal folds. Acute inflammation has been simulated by equation-based modeling [12, 18, 48, 76]. Kumar et al. [48] reported a three-variable ordinary differential equation model of acute inflammation. The model consists of a pathogen and two inflammatory mediators. In the model, an infectious agent triggered an early inflammatory mediator to kill the pathogen and later another inflammatory mediator was triggered to further exacerbate inflammation. The model was able to reproduce the health state and various negative states as observed in severely infected patients. Their simulation results also point out that the clinical presentations of severe infection arise from diverse underlying physiological states

and thus different therapeutic approaches are required.

Later, a study Reynolds et al. [76] reported a four-variable model of acute inflammation. This model is expanded from Kumar et al.'s [48] model by incorporating an anti-inflammatory mediator in the system. The anti-inflammatory mediator in the model has the functions of (1) inhibiting the build-up of pro-inflammatory mediator and the tissue damage, and (2) promoting further production of anti-inflammatory mediator. Their simulation results suggest that the rates of dynamic anti-inflammatory response may be modified therapeutically to yield optimal healing outcome after pathogenic infection.

Day et al. [18] used the same four-variable model of acute inflammatory response to explore the effects of repeated administration of endotoxin (bacterial lipopolysaccharide) on the system. Their model was capable to display the clinical scenarios of endotoxin tolerance and potentiation from a single parameter set with repeated endotoxin challenges. This finding suggests that various biological responses (endotoxin tolerance and potentiation) to repeated bacterial infection are dynamic manifestations of a unified acute inflammatory response.

Along the line of this research, a study reported the work of calibrating an acute inflammatory model to experimental data in mice with endotoxemia, surgical trauma or surgical trauma followed by hemorrhage [12]. The model was capable to predict a dose range of endotoxin leading to death of mice, despite being calibrated on data from non-lethal inflammatory insults. The most current model was calibrated to simulate acute inflammation in mice, swine and human in order to scale up the model to the whole-patient level [88].

Lastly, an equation-based model was developed to demonstrate the application of inflammation modeling for in silico clinical trial [13]. The model was set up to create a population of 1,000 synthetic "patients" with different clinical profiles. Then, these "patients" were submitted to one of three types of anti-sepsis therapy (anti-tumor necrosis factor neutralizing antibodies). Their simulation results suggest that anti-tumor necrosis factor therapy can be both effective and harmful to the patients, depending on the dose and duration of treatment as well as the patient's clinical profiles.

That said these models focused on the inflammation ensued from either bacterial infection or surgical trauma. In this study, we developed a new inflammation and healing model

to simulate early acute inflammatory responses (up to 24 hrs post phonotrauma) with input variables describing the vocal loading conditions and develop tools that will allow this model to effectively assimilate experimental data. We fitted the model against available experimental data and conducted parameter sensitivity analysis. We investigated the parameter space through metropolis sampling and parallel sampling method. Our goal is to develop a method for proposing a patient specific vocal exercise and rest schedule that optimizes tissue healing in cases of both acute and chronic phonotrauma.

3.2 EXPERIMENT, DATA AND ODE MODEL DEVELOPMENT

3.2.1 Experiment and Treatments

Data from 9 patients, six females (21-46 years old) and three males (21-49 years old), were collected in a study involving one hour loud phonation followed by a 4-hour treatment of three types: *voice rest*, *spontaneous speech* and *resonant voice* (roughly, classical singing) [53, 51]. One female subject received all three treatments, randomly ordered without replacement on different days by separated intervals ranging from 1-6 months. All the other 8 subjects participated in only one treatment. After the treatment at clinic, all subjects were discharged to home with instructions to continue their corresponding treatment in somewhat less intense cycles for resonant voice and spontaneous speech conditions. Laryngeal secretions are collected at 4 different times: at baseline, immediately after vocal loading, after the 4-hour clinic treatment and 24 hours post-baseline.

The detail time schedule of the whole experiment is shown in Figure 26. Patients arrived the clinic in the morning, and started laryngeal secretion baseline collection. Patients took 1 hour rest and then started the vocal loading, which has three consecutive cycles, each involving 15 minutes of loud phonation followed by 5 minutes of silence, for a total of 60 minutes. We assign the time when patients started vocal loading as time $t = 0$. Right after the vocal folding, time $t = 1$, laryngeal secretions are collected again and patients took 1 hour rest. Then the patients are randomly assigned to one of the three treatments: *voice*

rest, *spontaneous speech* or *resonant voice*. At time $t = 2.5$, each subject started their own treatment for 4 hours at the clinic. After the 4-hour treatment, laryngeal secretions were collected and patients took a rest. Then they went home at $t = 7.75$. At $10 < t < 14$, patients were instructed to continue their corresponding treatment at home and then went to bed after the treatment. They came back to the clinic next morning at $t = 22$.

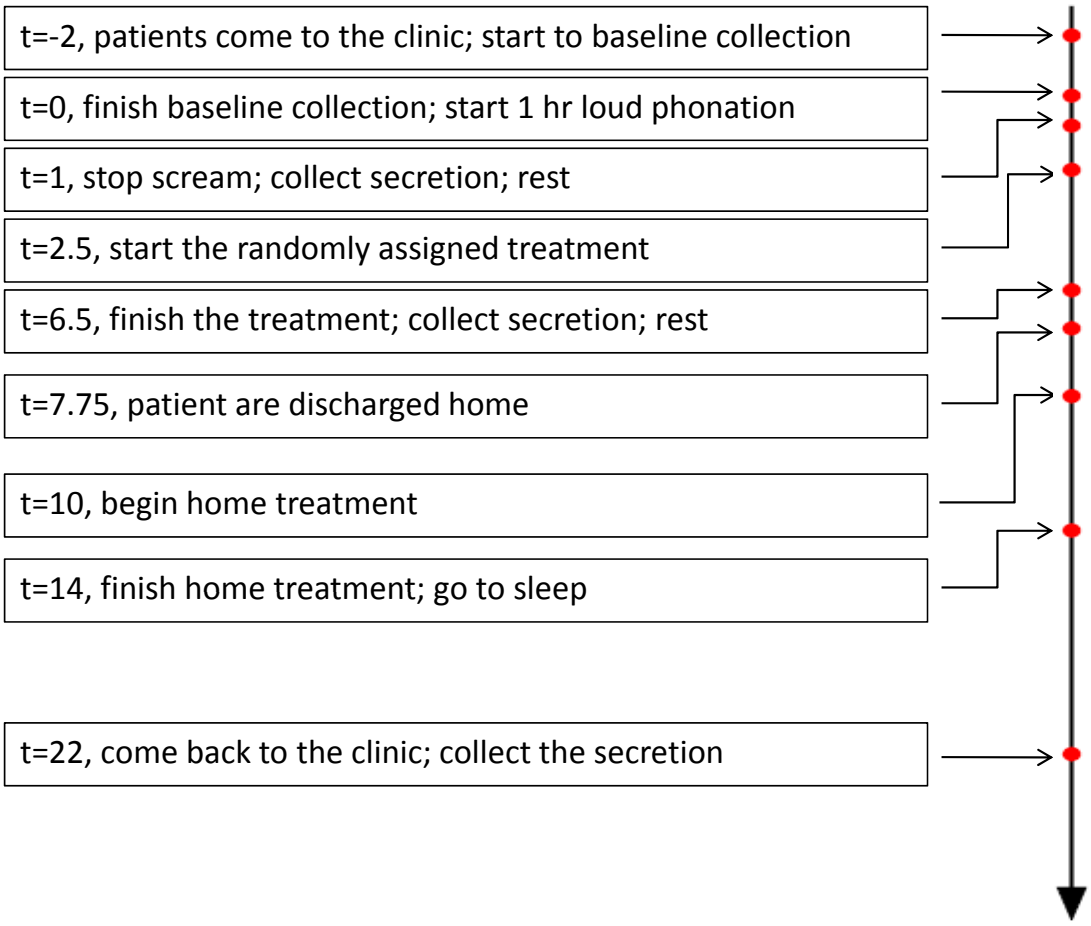

Figure 26: Treatment and data collection schedule. Loud phonation is performed during $0 < t < 1$, clinic treatment is during $2.5 < t < 6.5$, and home treatment is during $10 < t < 14$. Laryngeal secretion is collected at $t = 0, 1, 6.5, 22$.

The three treatment modalities can be considered on a continuum of tissue mobilization magnitude and intra-vocal fold impact stress magnitude. Voice rest has no mobilization or impact stress, spontaneous speech has normal to large magnitude mobilization but po-

tentially greater impact stress [73, 84], and resonant voice has normal to large magnitude but relatively low impact stress [73, 84]. It is believed that resonant voice treatment might result in a mostly anti-inflammatory response and therefore favors recovery from intense pro-inflammatory stimuli.

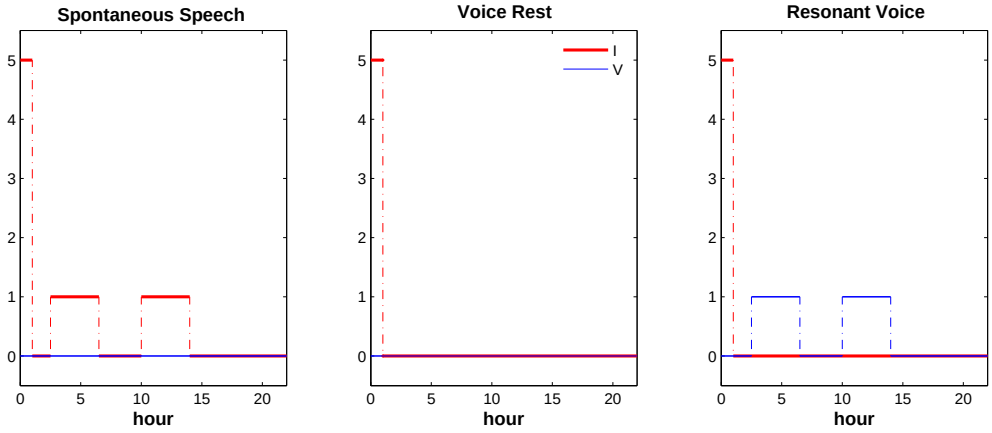

Figure 27: Vibration and impact stress loading values for the three treatments.

3.2.2 Data and ODE Model Development

Three subjects' data is excluded because of their thick secretions that are too viscous to pipette. One subject took all three treatments, thus we have nine valid treatment data trails. The spontaneous speech group has two subjects, the voice rest group has three, and the resonant voice has the left three. Figure 29 shows the average data of each treatment before exclusion of outliers. The first two collections are averaged over all valid data trails because the subjects are subject to the same conditions before they received their personal treatment. We have the measurements of TNF, IL-1β, IL-6, and IL-10.

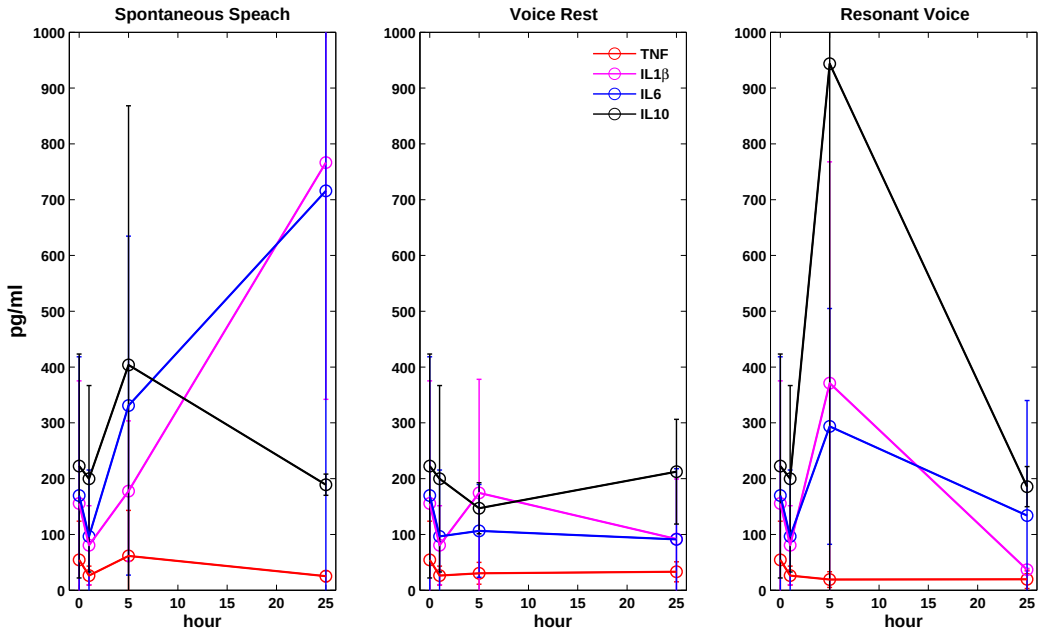

Figure 28: Average data of TNF, IL-1β, IL-6, and IL-10 in three treatment groups. The first two points are averaged over all the valid data set. The bars are the standard deviations.

In the early stage of the research, we developed the following model:

$$\Sigma = \frac{D + \alpha_{MIL1} IL1}{1 + \alpha_M (D + \alpha_{MIL1} IL1)} \tag{3.2.1}$$

$$\frac{dD}{dt} = M + \alpha_{DI1} I + \alpha_{DI2} DI - \mu_D D \tag{3.2.2}$$

$$\frac{dM}{dt} = \Sigma - \left(1 + \frac{\alpha_{MTGF1} TGF}{1 + \alpha_{MTGF2} TGF}\right) \left(\frac{IL10}{1 + \alpha_{MIL10} IL10}\right) \mu_M M \tag{3.2.3}$$

$$\frac{dIL1}{dt} = \frac{\alpha_{IL1M} M (1 + \alpha_{IL1s} \Sigma)}{1 + \alpha_{IL1IL10} IL10} - \mu_{IL1} IL1 \tag{3.2.4}$$

$$\frac{dIL6}{dt} = \frac{\alpha_{IL6M} M}{(1 + \alpha_{IL6IL10} IL10)} + \left(a_{IL6F} F + \frac{\alpha_{IL6F1} VF}{1 + \alpha_{IL6F2} V}\right) \left(1 + \frac{\alpha_{IL6IL11} IL1}{1 + \alpha_{IL6IL12} IL1}\right) - \mu_{IL6} IL6 \tag{3.2.5}$$

$$\frac{dIL10}{dt} = \frac{\alpha_{IL10M} M \left(1 + \alpha_{IL10V} V\right)}{1 + \alpha_{IL10IL10} IL10} - \mu_{IL10} IL10 \tag{3.2.6}$$

$$\frac{dTGF}{dt} = \alpha_{TGFM} M + \frac{\alpha_{TGFF} F}{1 + \alpha_{TGFIL1} IL1} - \mu_{TGF} TGF \tag{3.2.7}$$

$$\frac{dF}{dt} = \alpha_F \left(1 + \frac{\alpha_{FIL11} IL1}{1 + \alpha_{FIL12} IL1}\right)\left(1 + \frac{\alpha_{FTGF1} TGF}{1 + \alpha_{FTGF2} TGF}\right) - \mu_F F \tag{3.2.8}$$

$$\frac{dC}{dt} = \alpha_{CF} F \left(1 + \frac{\alpha_{CTGF1} TGF}{1 + \alpha_{CTGF2} TGF}\right) - \mu_C C \tag{3.2.9}$$

The model contains accounts for inflammation and healing processes. Two variables in the model describe the input vocal loading conditions: impact stress (**I**) and vibratory stress (**V**). The impact stress describes the magnitude of impact force exerted by one vocal fold on the other during strong phonation and is the primary source of damage in the system. [85]. The vibratory stress is generated by the vibration of vocal fold at any intensity and it simulates the production of cytokines by macrophages in the vocal fold [49]. Vibratory stress (**V**) and impact stress (**I**) are time-dependent variables according to the treatment schedule. The variable (**D**) represents damaged or dead cells. Macrophages (**M**) destroy the healthy cells and dominate in the early stage of inflammation. IL-1β (**IL1**) is pro-inflammatory cytokine [20] that is produced by the macrophages [19, 20] and is the most important cytokine for macrophage activation [59]. In the model, the damage and IL-1β form a pro-inflammatory signal Σ, which stimulates macrophage activation. IL-1β is inhibited by anti-inflammatory cytokine IL-10 (**IL10**) [59, 4, 79]. IL-6 (**IL6**) is another inflammatory mediator that is produced by macrophages [56] and fibroblasts in response to IL-1β [22]. and is inhibited by IL-10 [59, 79]. IL-10 is the main anti-inflammatory cytokine [1], which is also produced by macrophages [11], and inhibits its own production to a lesser degree than it inhibits other processes [17, 63]. TGF-β (**TGF**) is also an anti-inflammatory cytokine [1, 45], that is produced by macrophages and fibroblasts [45]. Exposure of vocal fold fibroblasts to IL-1β down-regulates TGF-β [54]. IL-10 and TGF-β deactivate macrophages [4, 5, 59]. Fibroblasts (**F**) are important sources of anti-inflammatories and their proliferation is stimulated by IL-1β and TGF-β [2, 23, 45]. Fibroblast population increases more than 1/3 under the influence of TGF-β [57]. Collagen (**C**) plays critical role in tissue development and repair [70] and it is mainly produced by fibroblast and TGF-β [7, 23, 55, 57].

The model (3.2.1)-(3.2.9) contains 8 variables and 40 parameters. Ideally, we would like to estimate the parameters and expect the model effectively assimilate all the experimental data. The difficulty here is that we only have valid data of IL-1β, IL-6 and IL-10 at 4 time

points. In view of the limited available data, our model has too many degrees of freedom to be fitted.

In order to avoid overfitting, we reduced our model to a four-variable model. Since variables IL6, TGF, F and C are essentially decoupled from the others and not supported by data, we eliminate them from the model. The reduced ODE model is shown here

$$\Sigma = \frac{\alpha_{M1}(D + \alpha_{MIL1} IL1)}{1 + \alpha_{M2}(D + \alpha_{MIL1} IL1)} \tag{3.2.10}$$

$$\frac{dD}{dt} = \alpha_{DM} M + \alpha_{DI1} I + \alpha_{DI2} DI - \mu_D D \tag{3.2.11}$$

$$\frac{dM}{dt} = \Sigma - \frac{IL10}{1 + \alpha_{MIL10} IL10} \mu_M M \tag{3.2.12}$$

$$\frac{dIL1}{dt} = \frac{\alpha_{IL1M} M(1 + \alpha_{IL1s}\Sigma)(1 + \alpha_{IL1V} V)}{1 + \alpha_{IL1IL10} IL10} - \mu_{IL1} IL1 \tag{3.2.13}$$

$$\frac{dIL10}{dt} = \frac{\alpha_{IL10M} M(1 + \alpha_{IL10V} V)}{1 + \alpha_{IL10IL10} IL10} - \mu_{IL10} IL10 \tag{3.2.14}$$

The reduced model was capable of explaining the data within the standard deviation. We have explored the whole parameter space using the parallel sampling method, finished the sensitivity analysis and parameter reduction, and fitted the reduced parameter set to the data from individual patients. Although results were numerically acceptable, we were concerned with the extensive variability observed between individuals. Should we just simply take average data (naive pooling) of them and fit them despite the huge standard deviations? For example, the peak of IL-10 in resonant voice group is pushed by a singe point, which is out the plot because of its large value (> 2400). We also observed that some of the patients have a high baseline and should be considered as pre-inflamed. It did not appear suitable to take average over these pre-inflamed patients together with others.

To address these concerns, we went back and carefully reviewed the experimental data. Each secretion collection of a single patient is measured twice. The primary measurements show that two patients in the voice rest group have a high initial value of IL-1β, between 500 ng/ml to 1100 ng/ml. These two should be separated as an additional group called voice rest pre-inflamed group. Thus, we have three subjects in the spontaneous speech group, two

in the voice rest group, two in the voice rest pre-inflamed group and three in the resonant group. All the primary collection measures are depicted in Figure 29.

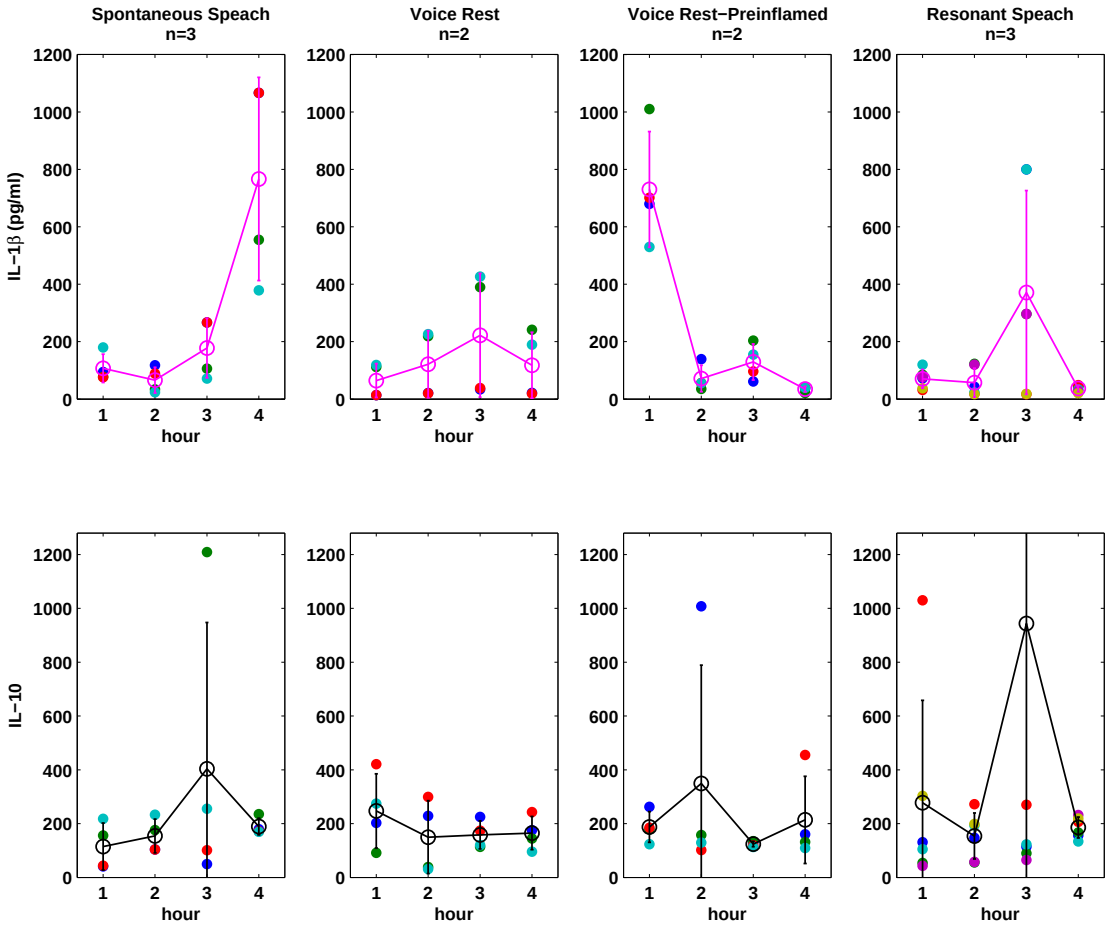

Figure 29: Valid primary collection measurements. n is the number of subjects in each group. Four time points - baseline, after 1 hour screaming, after 4 hour clinic treatment and 22 hours post screaming. The lines and bars represent the average and standard deviation of each group. The highest point of IL-1β in spontaneous group is overlapped by two measurements and so does that in resonant voice group. The lines are the average curves and bars are standard deviations.

In Figure 29, every dot represents the measurement of each subject's primary collection at 4 time points. The highest point of IL-1β in the spontaneous group is overlapped by two measurements and as is the case in the resonant voice group. In the spontaneous speech treatment group, one IL-10 point (> 1200) is significantly higher than the other

measurements (< 400) at the same time point. So we treated this measurement as an outlier. IL-10 data in the voice rest pre-inflamed group has an outlier with value greater than 1000 while all the other data at that time is below 200. In the data of IL-10 in the resonant group, there is one point at $t = 6.5$ that not visible in the plot because it is greater than 5,000 and one of its baseline is greater than 1000 while all the others are below 350. These two measurements should also be treated as outliers.

Figure 30 plots all the primary measurements without outliers.

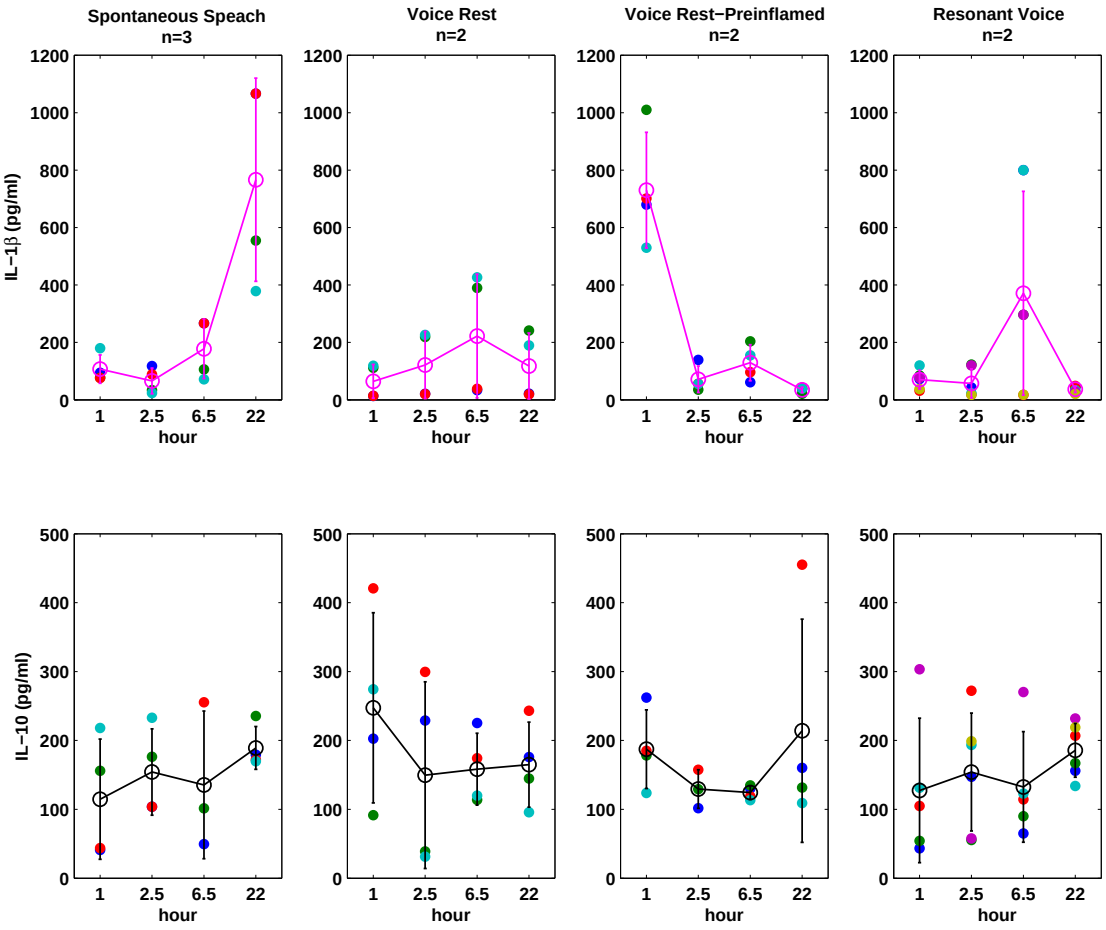

Figure 30: Valid primary collection measurements with outliers eliminated. n is the number of subjects in each group. Four time points - baseline, after 1 hour screaming, after 4 hour clinic treatment and 22 hours post screaming. The lines and bars represent the average and standard derivation of each group.

After excluding outliers, the large standard derivations are still troublesome. The average

baseline in voice rest treatment group is almost double of those of the other two groups. But we expect the same baselines for the three groups since we are dealing with the average data. Therefore we need to normalize each data set individually. Each individual baseline has two measurements, so the normalized factor is the average of the two. We shall fit the reduced model to all treatment groups except the pre-inflamed group.

In Figure 30, before normalization, IL-1β peaks higher and later in spontaneous speech group than the other two groups and it decreases faster with resonant voice than with voice rest. The trajectories of IL-10 in the spontaneous speech group and resonant voice group are almost the same and that in voice rest group never peaks - it decreases first and remains flat all the time. In Figure 31, after normalization, we can see that IL-1β trajectories in each group behave similarly as before, but the trajectory of IL-10 in spontaneous speech peaks higher than in resonant voice at $t = 22$, and that in resonant voice IL-10 also has a small peak. The normalized data set is more reasonable in a biological sense. The final normalized data are in Table 1, 2, and 3.

Resonant has higher magnitude of vibration stress loading and lower magnitude of impact stress loading than spontaneous. The IL-10 equation (3.2.14) mainly depends on IL-10 itself and the vibration stress. It is hard for (3.2.14) to predict similar IL-10 trajectories of spontaneous and resonant groups at the same time because of their different input vibration stresses. Thus we updated our model again.

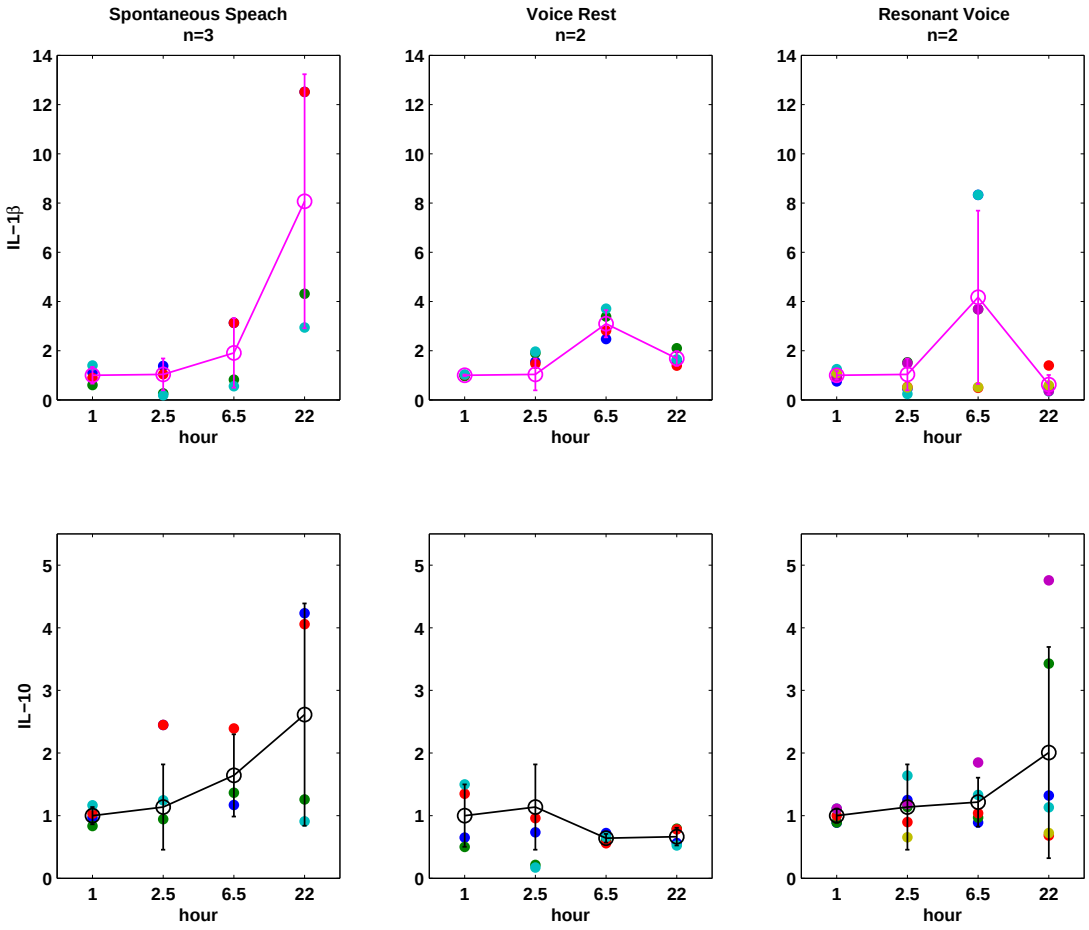

Figure 31: Normalized data excluding the pre-inflamed patients and outliers. n is the number of subjects in each group. The lines and bars represent the averages and standard derivations. The second point is averaged over all valid data.

Time (hour)	P (average)	P (stand deviation)	A (average)	A (stand deviation)
0	1.00	0.34	1.00	0.14
1	1.03	0.65	1.14	0.68
6.5	1.91	1.41	1.64	0.66
22	8.07	5.16	2.61	1.78

Table 1: Normalized data for the spontaneous speech group

Time (hour)	P (average)	P (stand deviation)	A (average)	A (stand deviation)
0	1.00	0.03	1.00	0.50
1	1.03	0.65	1.14	0.68
6.5	3.09	0.55	0.64	0.07
22	1.68	0.30	0.66	0.14

Table 2: Normalized data for the voice rest group

Time (hour)	P (average)	P (stand deviation)	A (average)	A (stand deviation)
0	1.00	0.17	1.00	0.11
1	1.03	0.65	1.14	0.68
6.5	4.17	3.52	1.22	0.39
22	0.62	0.39	2.01	1.69

Table 3: Normalized data for the resonant voice group

3.3 ODE MODEL

Because of the limited available data, we designed a reduced model that contains four state variables representing the amount of tissue damage (**D**), activated inflammatory cells (**M**), pro-inflammatory cytokines (**P**) and anti-inflammatory cytokines (**A**). The inflammatory cells are represent macrophages, neutrophils or natural killer cells, pro-inflammatory cytokines represent, for example, TNF-α or IL-1β, and the anti-inflammatory cytokines represent IL-10. The influence diagram for the model is shown in Figure 32 .

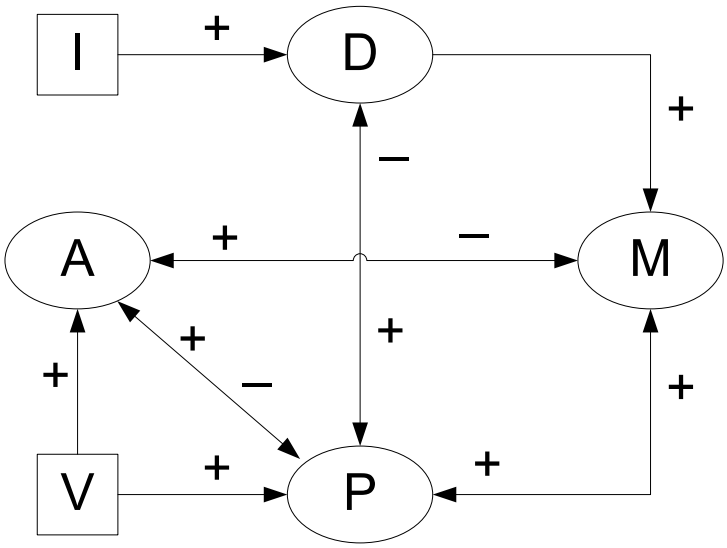

Figure 32: Model schema that illustrates the interactions between the four variables. The plus indicates positive feedback and minus indicate negative feedback.

The model equations are here:

$$\frac{dD}{dt} = \alpha_{DI} I - \frac{P^2}{\alpha_{DP}^2 + P^2}\mu_D D \tag{3.3.1}$$

$$\frac{dM}{dt} = \alpha_{MD}(D + \beta P) - \frac{\varepsilon + A}{1 + \alpha_{MA} A}\mu_M M \tag{3.3.2}$$

$$\frac{dP}{dt} = \frac{\alpha_{PM} M(1 + \alpha_{PV} V)}{1 + \alpha_A A^2}(1 + \alpha_{PD} D) - \mu_P P \tag{3.3.3}$$

$$\frac{dA}{dt} = \frac{\alpha_{AM} M(1 + \alpha_{PV} V)}{1 + (\alpha_A/\gamma)A^2}\frac{P^2}{\alpha_{AP}^2 + P^2} - \mu_A A \tag{3.3.4}$$

Tissue damage (**D**) is primarily created by impact stress (**I**) in the vocal fold [37, 38]. Experimental data show that the inflammatory response is mild. Therefore we neglect secondary damage creation due to the presence of neutrophils that is prominent in some published models of sepsis [48]. On the contrary, we assume that mild inflammation has a beneficial effect of acceleration of healing of the damage, represented by a saturating term. D together with pro-inflammatory cytokines (**P**) stimulates the activation of macrophages (**M**) [59, 4, 79]. D also stimulates the production of P by M [19, 20]. M produce anti-inflammatory cytokine (**A**) [11], that result in M deactivation [4, 5, 59] and inhibition of the

M production of P [4, 79]. A also inhibits its own production a lesser degree than it inhibits the production of P [17, 63], so the γ in equation (3.3.4) should be greater than one. The production of A is stimulated by the presence of P through a sigmoidal function.

The prescribed dependence of V and I on time is determined by the experimental loading schedule for the insult and the various treatments. In particular, we assume the loaded impact stress I during the first hour screaming is 5 and the vibration stress V during that period is 0. We take the values of vibratory stress during the spontaneous speech, voice rest and resonant voice, respectively, to be 0, 0, and 1. Similarly, the values of impact stress during those treatments were taken to be 1, 0 and 0 [85, 53]. The treatment duration is between time 2.5h to 6.5h at the clinic and between 10h and 14h at patients' home. The values of model parameters were kept the same for all treatments.

3.3.1 Stability Analysis

In the absence of external input (i.e., if V and I are zero) and for small values of ε, the ODE system has an unstable equilibrium at $(D, M, P, A) = (0, 0, 0, 0)$ and a stable equilibrium at $(D, M, P, A) = (0, \bar{M}, \bar{P}, \bar{A})$, for which $\bar{P}$ and $\bar{A}$ are close to 1.

Steady state values of the system are solutions of the following system of equations:

$$0 = \frac{P^2}{\alpha_{DP}^2 + P^2}\mu_D D \tag{3.3.5}$$

$$0 = \alpha_{MD}(D + \beta P) - \frac{\varepsilon + A}{1 + \alpha_{MA}A}\mu_M M \tag{3.3.6}$$

$$0 = \frac{\alpha_{PM}M}{1 + \alpha_A A^2}(1 + \alpha_{PD}D) - \mu_P P \tag{3.3.7}$$

$$0 = \frac{\alpha_{AM}M}{1 + (\alpha_A/\gamma)A^2}\frac{P^2}{\alpha_{AP}^2 + P^2} - \mu_A A \tag{3.3.8}$$

Equation (3.3.5) implies that $P = 0$ or $D = 0$. If $P = 0$, then the remaining equations imply that $M = 0$, $A = 0$, and $D = 0$ and hence there is a zero equilibrium of the system $(D, M, P, A) = (0, 0, 0, 0)$. This equilibrium is unstable.

If $D = 0$, then M, P and A obey the following system:

$$0 = \alpha_{MD}\beta P - \frac{\varepsilon + A}{1 + \alpha_{MA}A}\mu_M M \tag{3.3.9}$$

$$0 = \frac{\alpha_{PM} M}{1+\alpha_A A^2} - \mu_P P \tag{3.3.10}$$

$$0 = \frac{\alpha_{AM} M}{1+(\alpha_A/\gamma)A^2} \frac{P^2}{\alpha_{AP}^2 + P^2} - \mu_A A \tag{3.3.11}$$

Equations (3.3.9) and (3.3.10) imply that the equilibrium value of A is equal to a root of the following equation.

$$(\varepsilon + A)(1+\alpha_A A^2) - \frac{\alpha_{MD}\beta\alpha_{PM}}{\mu_M \mu_P}(1+\alpha_{MA}A) = 0 \tag{3.3.12}$$

or

$$\alpha_A A^3 + \varepsilon\alpha_A A^2 + \left(1 - \frac{\alpha_{MD}\beta\alpha_{PM}\alpha_{MA}}{\mu_M\mu_P}\right) A + \left(\varepsilon - \frac{\alpha_{MD}\beta\alpha_{PM}}{\mu_M\mu_P}\right) = 0 \tag{3.3.13}$$

According to the Routh criterion the number of non-negative roots is equal to the number of sign changes in the left column of the array

$$\begin{pmatrix} \alpha_A & 1 - \frac{\alpha_{MD}\beta\alpha_{PM}\alpha_{MA}}{\mu_M\mu_P} \\ \varepsilon\alpha_A & \varepsilon - \frac{\alpha_{MD}\beta\alpha_{PM}}{\mu_M\mu_P} \\ \frac{\alpha_A\alpha_{MD}\beta\alpha_{PM}}{\mu_M\mu_P}\left(\frac{1-\varepsilon\alpha_A\alpha_{MA}}{\varepsilon\alpha_A}\right) & 0 \\ \frac{\alpha_A\alpha_{MD}\beta\alpha_{PM}}{\mu_M\mu_P}\left(\frac{1-\varepsilon\alpha_A\alpha_{MA}}{\varepsilon\alpha_A}\right)\left(\varepsilon - \frac{\alpha_{MD}\beta\alpha_{PM}}{\mu_M\mu_P}\right) & 0 \end{pmatrix} \tag{3.3.14}$$

If ε is small enough, more precisely, if $\varepsilon < min\left(\frac{1}{\alpha_A\alpha_{MA}}, \frac{\alpha_{MD}\beta\alpha_{PM}}{\mu_M\mu_P}\right)$, then the array has precisely one change in sign, as the first three entries are positive and the last is negative. Hence there is one positive solution for A. Equations (3.3.9) and (3.3.10) imply that M and P have the same sign in equilibrium and (3.3.11) implies that M and A have the same sign. Thus, for small ε there is a single positive equilibrium $(D, M, P, A) = (0, \bar{M}, \bar{P}, \bar{A})$, which is linearly stable.

3.3.2 Parameter Estimation

Parameter values were estimated using both traditional maximum likelihood optimization and also using a novel method of ensemble modeling. Traditional parameter fitting methodology is based on the assumption that all experimental subjects can be described by the same set of parameters and that random noise leads to differences between the measured values of state variables. By taking averages over multiple subjects, one can recover the true value of the variable and then optimize the parameters of the model so that the model predicts the true values.

3.3.2.1 Objective Function in the Optimization The first step of model fitting is to optimize parameter values over all three treatment scenarios simultaneously. The objective function F reflects the difference between model prediction and the available data. Two versions were considered - one in which the mean square differences were weighted by the variance of the data and another in which all data were taken with equal weights. Calibration with the first choice of function led to an unresponsive model with flat configurations, which is why we gave preference to the second choice:

The objective function used in the optimization is constructed in the following way:

$$\begin{aligned} F &= \sum_{j=1}^{3}\sum_{i=1}^{4}\left(y_{ij}^{P}-\widetilde{y}_{ij}^{P}\right)^{2}+\sum_{j=1}^{3}\sum_{i=1}^{4}\left(y_{ij}^{A}-\widetilde{y}_{ij}^{A}\right)^{2}+\sum_{j=1}^{3}\left(M_{0,j}-M_{0}\right)^{2} \\ &\quad +\sum_{j=1}^{3}\left(D_{\max,j}-\overline{D}\right)^{2}+\sum_{j=1}^{3}\left(D_{end,j}\right)^{2} \end{aligned} \tag{3.3.15}$$

In the objective function (3.3.15) the index j represents the summation over the data in three treatments and the index i represents summation over the four time points. y_i^P and y_i^A are the simulated data computed from the trajectory of the model, while $\widetilde{y}_i^P$ and $\widetilde{y}_i^A$ are observed data from the experiment - the normalized data that we discussed in last section. Since we have normalized the data by individual baseline, the starting values of P and A are 1. Similarly, M also has an initial baseline to start with, which we define as M_0 with value 1. The initial value of damage is 0. As additional heuristic criteria we have included the final value of the damage $D_{end,j}$ (assumed to be zero as in a successful healing process)

and the maximum value of the damage $D_{max,j}$ to be an arbitrary value $\overline{D}$ in order to fix the scale of that variable. The value $\overline{D}$ is assumed to be the same over different treatments as the treatments are not expected to add additional damage to the system after the first hour.

To find the set of parameters that minimize (3.3.15) we use an iterative optimization algorithm. For each set of parameters, we find the corresponding equilibrium value and then use it as the initial value for integration of the system. The resulting trajectory is the used to compute y_i^P and y_i^A. This way we can make sure that the optimal model has an equilibrium point close to the homeostatic level. The data used for fitting are shown in Table 1, 2, and 3.

3.3.2.2 Ensemble Model Ensemble modeling replaces a traditional, uniquely parameterized model with an ensemble of models that vary in parameter values and possibly the model structure. Each model in the ensemble is assigned a probability value that correlates with the likelihood that the model agrees with given data and heuristic constraints. In this way the ensemble of models summarizes uncertainties in parameter values, sensitivity of the model, and the completeness of the data in a single distribution.

One way to interpret the ensemble model is by treating it as a representation of population heterogeneity of parameter values, i.e., the fact that parameters may be distinct for different individuals in the population.

The main benefit of ensemble model is that any predictions made using the model is probabilistic, resulting from all models compatible with the data but with larger weight coming from more accurate models. Therefore, for each such a prediction one can determine the interval of confidence.

The idea of ensemble modeling was first introduced in weather prediction but has recently been applied also to biological models. The most common ensemble modeling method constructs the ensemble by fixing the structure of the model (i.e., the set of differential equations and definitions of the parameters) and then computes the probability density $\rho(\alpha)$ of a parameter set α with the likelihood $L(\alpha)$ of observing the data with that model. The likelihood function depends on the amount and quality of experimental data. The probability density is computed using Bayesian approach as $\rho(\alpha) = Q^{-1}L(\alpha)\theta(\alpha)$, where $Q = \int L(\alpha)\theta(\alpha)d\alpha$ is

a normalizing constant and $\theta(\alpha)$ is the *prior* distribution containing information about the range of parameter values that is biologically acceptable. Here we sample Gibbs-Boltzman distribution $\rho(\alpha) = \exp(-\beta F(\alpha))$, where $F(\alpha)$ is the objective function corresponding to the parameter set α and β is the reciprocal of temperature value T. We choose $\theta(\alpha)$ for each parameter to be a uniform distribution roughly between 1/20 and 20 times the baseline value. Table 4 shows the baseline, prior range, and the average and standard deviation for each parameter. We choose α_i to represent the logarithm of the value of the corresponding parameter (i.e., $\alpha_1 = \log \alpha_{DI}$, $\alpha_2 = \log \mu_D$) rather than the parameter itself for several reasons: (i) trajectories of a system of ODEs are sensitive to relative, rather than absolute changes in parameters, (ii) the parameters are always positive and (iii) we wish to avoid biasing the sample toward large values of the parameters. This choice does not affect the principles of the ensemble modeling, only the sampling metric on the parameter space.

More discussion about ensemble methods can be found in Hancioglu's PhD thesis [39].

Metropolis Sampling

The easiest way to represent the distribution $\rho(\alpha)$ is by a sample of points α^1,..., α^M distributed according to $\rho(\alpha)$. The sample can be used to estimate the ensemble average of any trajectory-dependent quantity H as

$$< H > \cong M^{-1} \sum_{i=1}^{M} H(\mathbf{x}(t; \alpha^i, \mathbf{x}_0)) \tag{3.3.16}$$

The percentile value $P_X(H)$ of any trajectory-dependent quantity H can be found as the smallest number that is larger than $X\%$ of values $H(\mathbf{x}(t; \alpha^i, \mathbf{x}_0)), i = 1, ..., M$.

The Metropolis-Hastings Markov Chain (MHMC) sampling algorithm is a Monte Carlo method for finding a sample α^1,..., α^M for any distribution defined as $\rho(\alpha) = exp(-\beta E(\alpha))$ [62, 41]. The original MHMC method samples the distribution at a fixed value of β by proposing, at every step k, a random perturbation $\bar{\alpha}$ of the current parameter set α^k and accepting this perturbed set as the next set in the sample, i.e., α^{k+1}, with the probability $P = min\{1, exp\left(-\beta\left(E(\bar{\alpha}) - E(\alpha^k)\right)\right)\}$. If $\bar{\alpha}$ is rejected then α^{k+1} is set equal to α^k. In other words, if the proposed move is associated with a reduction of the objective function, thus a likelihood ratio > 1, the move is accepted. Yet, if the likelihood ratio is < 1, there is still a probability < 1 that the move is accepted, allowing the possibility of the chain to

escape local minima. The resulting sequence α^k is a Markov Chain that has a stationary probability distribution that converges to $\rho(\alpha)$ in the limit of large k.

In our case, we define the objective function (3.3.15) as an energy function E and sample Gibbs-Boltzmann distribution through $\rho(\alpha) = exp\,(-\beta E(\alpha))$. The algorithm starts at some random initial guess of parameters, and then let these parameters jump to its neighborhood. Calculate the energy function of the two parameters set. If the new energy is smaller than the old one, then the new parameter set is accepted with probability 1, otherwise it is accepted with a probability, which is equal to the log likelihood ratio. A random number is generated from the uniform distribution on zero to one each step and compared with the log likelihood ratio. If the number is greater than that ratio, then the move is accepted, otherwise, it is rejected. When the number of sampling steps goes to infinity, the probability of parameters values converges to an equilibrium probability distribution, which is independent of the initial parameter values [27].

The algorithm is below [39]:

- Construct the energy function E.
- Start at some random initial guess of parameter set α_1 and calculate the value of the energy function E as E_1 at α_1.
- Generate a new parameter set $\bar{\alpha}$ through the formula $\bar{\alpha} = \alpha_1 exp\,(-\varepsilon r_n)$. ε is a constant associated with the temperature $T = 1/\beta$, r_n is a rand number generated by the standard normal distribution.
- Calculate the energy E_b at $\bar{\alpha}$ and accept $\bar{\alpha}$ with a probability of $P = min\,(1, exp\,(-\beta(E_b - E_1)))$.
- If Y is accepted, then assign $\alpha_2 = \bar{\alpha}$ and $E_2 = E_b$, else $\alpha_2 = \alpha_1$ and $E_2 = E_1$.
- Repeat from step 3. Adjust the ε value to get the acceptance rate around 30%.

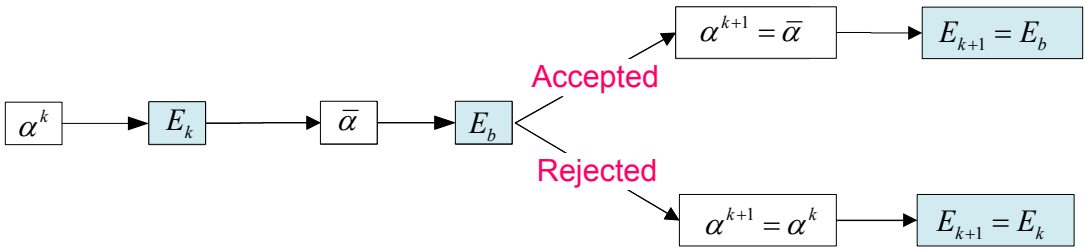

Figure 33: Schematic representation of the Metropolis sampling algorithm.

Parallel Tempering

Unlike Metropolis sampling, which is consisted of a single Markov chain at one particular temperature value, parallel tempering algorithm generates multiple versions $\alpha^{k,i}$ of the Markov Chain evaluated at different temperature values (i.e., different β). At high temperature (low β), the parameters have a bigger step size so that they can explore a larger region of the phase space. At low temperature, the parameters are more precisely sampled in a local region [25]. Every once in a while the parameter sets $\alpha^{k,i}$ and $\alpha^{k,i+1}$ in two neighboring values β_i and β_{i+1} are swapped with probability $P = min\{1, exp([\beta_i - \beta_{i+1}][E(\alpha^{k,i}) - E(\alpha^{k,i+1})])\}$. Therefore, using parallel tempering the algorithm can explore larger phase space and converge faster to a stationary probability distribution by improving chain mixing.

We construct 4 replicas with β_1, β_2, β_3 and β_4 equals to 1, 0.5, 0.25 and 0.125, respectively. Optimal convergence of parallel tempering MHMC is achieved if the acceptance ratio (i.e, the ratio of accepted to proposed sets) has a specified value. Rathore et al. suggested that the optimal acceptance ration of each replica is 20% to yield the best performance, Kone et al. suggested this ratio to be 23% [25], while Feller W. suggested it in the range of 30% to 40% [27]. In our simulation, the accepted ratios of each replica are 30.2%, 30.6%, 30.9% and 31.3%. The ε values, a measure of the proposed steps in the procedure are chosen as 0.071, 0.108, 0.162 and 0.249 of each replica. The chain with $\beta = 1$ provides the Boltzmann sample. The swapping acceptance rate between replica 1 & 2, 2 & 3, 3 & 4 are 26.0%, 28.3%, 37.5%.

The algorithm we use is as follows:

- Construct 4 replicas by using Metropolis sampling based on different temperature values.
- Run the basic Metropolis algorithm for each replica five steps. Five is chosen to give a reasonable swapping rate, therefore to provide a good mixing.
- Perform swaps between adjacent replicas with a probability of $P = min\{1, exp([\beta_i - \beta_{i+1}][E(\alpha^{k,i}) - E(\alpha^{k,i+1})])\}$, where $\beta_i > \beta_{i+1}$
- Repeat from step 2.

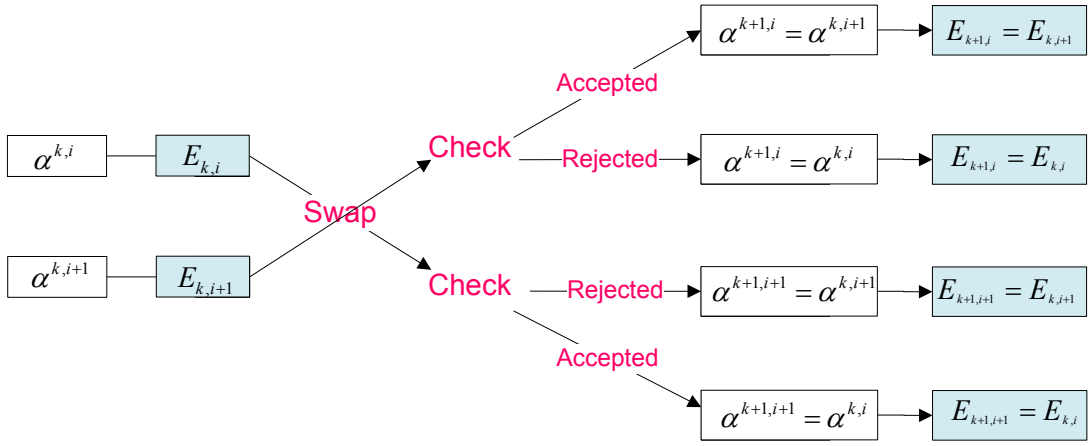

Figure 34: Schematic representation of the Parallel-Tempering algorithm.

3.3.3 Sensitivity Analysis and Parameter Reduction

The sensitivity analysis reveals the robustness of the model by exploring to which parameters the trajectory of system is most sensitive. One can use the results to reduce the number of parameters in the model.

The model (3.3.1) $\sim$ (3.3.4) has 4 equations with 4 states (D, M, P, A) and 17 parameters (θ). Since we only have data of P and A, so we only consider sensitivities at those two equations. The 2×17 sensitivity matrix $S = (s_{ij})$ is defined below. The sensitivity coefficients s_{ij} are calculated from the difference of nominal and perturbed solutions [16].

$$s_{ij}(t) = \frac{\partial x_i(t)}{\partial \theta_j} = \frac{x_i(\theta_j + \Delta\theta_j, t) - x_i(\theta_j, t)}{\Delta\theta_j} \tag{3.3.17}$$

Here $i \in [1, 2]$ and $j \in [1, 17]$. Since the states and parameters have different magnitudes, in order to facilitate the comparison of parameter sensitivities, we normalized s_{ij} in the following way [16]:

$$\bar{s}_{ij} = \frac{\partial x_i(t)}{\partial \theta_j} \frac{\theta_j}{x_i} \tag{3.3.18}$$

Calculate the normalized sensitivity matrix $\bar{S}(t)$ at 4 data points. The matrix G is constructed in the following way:

$$G = \begin{bmatrix} \bar{S}(t_1) \\ \bar{S}(t_2) \\ \bar{S}(t_3) \\ \bar{S}(t_4) \end{bmatrix} \tag{3.3.19}$$

We calculate the G matrices over 2,000 parameter samples evenly chosen from samples generated by parallel tempering and take the average over all these G matrices to get a new matrix $\overline{G} = \sum_{i=1}^{n} G_i$. Then calculate the covariance matrix C, where $C(i, j)$ is the covariance of the ith and jth column of $\bar{G}$. By sorting the diagonal elements of matrix C, we can find the most to least sensitive parameters. Since some of the parameters are highly correlated, we need to check the covariance coefficient matrix R, which $R(i, j) = \frac{C(i,j)}{\sqrt{C(i,i)C(i,j)}}$. If some pair of the parameters are highly correlated with each other, then we only need to pick one of them. Here is the algorithm to pick the reduced parameter set.

- Calculate the sensitivity matrices G over the parameter sets got from parallel tempering.
- Calculate the averaged sensitivity matrix $\overline{G} = \sum_{i=1}^{n} G_i$.
- Calculate the covariance matrix C of $\bar{G}$ and sort the diagonal elements of C. Then we have the sensitive parameters in order.
- Start from the most sensitive parameter i, and check its covariance coefficients $R(i, j)$ with others. Exclude the parameters which are highly correlated with i. And then repeat this procedure to the next sensitive parameter if it has not been excluded yet.

3.3.4 Individual Fit

In the sensitivity analysis, we determine parameters to which the model is the most sensitive. We perform parameter fitting for individuals by minimize the objective function

over the sensitive parameters while fixing the other parameters at the average values of the marginal distributions obtained from parallel sampling.

3.3.5 ODE Model Results

All the numerical results are simulated in MATLAB 7.5.0 (R2007b). The collected data are shown in Figure 30 after elimination of outliers (defined as isolated values outside of 3 standard deviations for each data point). The lines and bars represent the average and standard derivation of each group.

From the data one can see that IL-10 is quite flat during the observation period of any group. Since the baseline values varied over an order of magnitude, we re-normalized the data by dividing data for each subject by that subjects average base-line values (at the first time point). Normalized data are shown in Figure 31. In that figure, the average at the first and second time points is averaged over all patient data at that time point because all patients to the same conditions up to the point, no matter which treatment group they belong to.

There are statistically significant differences between the trajectory of spontaneous speech and the trajectories of the other two treatments. Much less difference can be observed between voice rest and resonant voice treatment. Spontaneous speech appears to lead to a high level of inflammation at the end of the 24hr period. Voice rest and resonant voice treatments have IL1-β levels that peak in the middle of the treatment period and then return to the baseline by 24hr. The peak for resonant voice is slightly larger but with larger variance as well. IL-10 levels rise during treatment phase for both spontaneous speech and resonant voice treatments but not for voice rest treatment.

To explain the data we hypothesize that spontaneous speech increases damage in the system by producing moderate levels of impact stress. The additional damage leads to increased levels of pro-inflammatories for that treatment. We further hypothesize that resonant voice produces large amount of vibratory stress that increases the production of both inflammatory cytokines by macrophages. We have utilized these hypotheses in the construction of the model and then set out to validate those hypotheses by fitting the model to the data.

The optimal parameter fit, based on minimization of the objective function (3.3.15), can be seen in Figure 35. The last column of Table 4 shows the parameter values represented in that figure. The trajectory of the system fits most of the data points. The exception is the level of IL-10 in the voice-rest treatment case that is predicted 3 times larger than the experimental values. The trajectory for macrophages is compatible with mild inflammation, resulting in up to 3 fold activation of macrophages (for spontaneous speech). The trajectory for damage is almost identical for voice rest and resonant voice treatments.

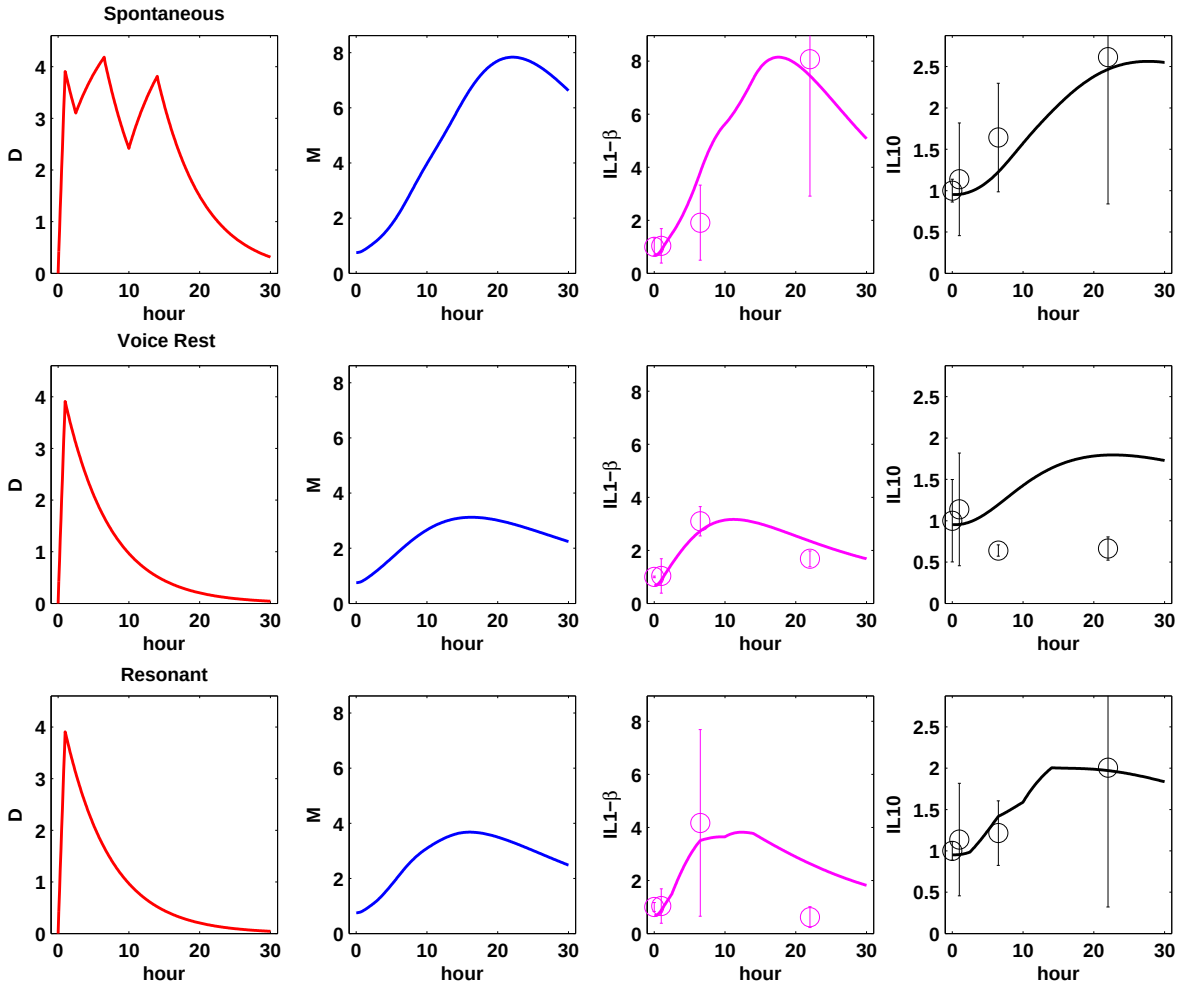

Figure 35: Optimal parameter fit for the model. The parameter values are showed in the last column of Table 4.

3.3.5.1 Ensemble Method Table 4 shows the baseline, prior range, average and standard deviation for each parameter.

No	Parameter	Baseline	Range	Average	SD	Sample
1	α_{DI}	0.86	$0.043 \sim 17.2$	0.7631	0.1172	0.8420
2	α_{DP}	0.53	$0.0265 \sim 10.6$	0.3194	0.3974	0.1338
3	μ_D	0.18	$0.009 \sim 3.6$	0.133	0.0681	0.1555
4	α_{MD}	0.05	$0.0025 \sim 1$	0.0566	0.0943	0.0339
5	β	1.8	$0.01 \sim 10$	2.5262	2.5955	5.4937
6	ε	0.1	$0.005 \sim 1$	0.1239	0.1537	0.0069
7	α_{MA}	3.4	$0.17 \sim 68$	11.7458	15.6568	6.6537
8	μ_M	0.3	$0.015 \sim 6$	0.3912	0.7383	1.2744
9	α_{PM}	0.17	$0.085 \sim 3.4$	0.4534	0.5541	0.1988
10	α_{PD}	1.3	$0.065 \sim 26$	4.8495	4.3919	1.6785
11	α_{PV}	1	$0.05 \sim 20$	0.8571	0.8845	0.5789
12	α_A	1.5	$0.075 \sim 30$	8.0904	8.2102	1.6323
13	μ_P	0.14	$0.025 \sim 0.5$	0.1037	0.0628	0.0896
14	α_{AM}	0.1	$0.025 \sim 10$	1.7211	2.0583	0.1735
15	γ	4	$1 \sim 10$	2.9009	2.0393	1.7598
16	α_{AP}	2.7	$0.2 \sim 4$	0.4386	0.2520	0.2022
17	μ_A	0.1	$0.05 \sim 1$	0.2909	0.2524	0.0682

Table 4: Baseline, prior range, average and standard deviation of each parameter

We have performed 4 distinct runs starting from randomly initial parameter values by the Metropolis Monte Carlo algorithm. The length of each run is 320,000, 320,000, 320,000 and 90,000 after discarding the initial points (burn-in-period) from each run. In order to lower the autocorrelation, the samples are stored every 5th point. The convergence of each run is tested using Gelman and Rubin diagnostic [28], which calculates a potential scale reduction factor associated with each parameter. That factor values in our model are all below 1.01 which is within the confidence interval.

We have combined all stored samples of 210,000 points for further analysis. Marginal

parameter distributions are shown in Figure 36. 14 parameters have peaks within the prior distribution range, while the remaining distributions are skewed.

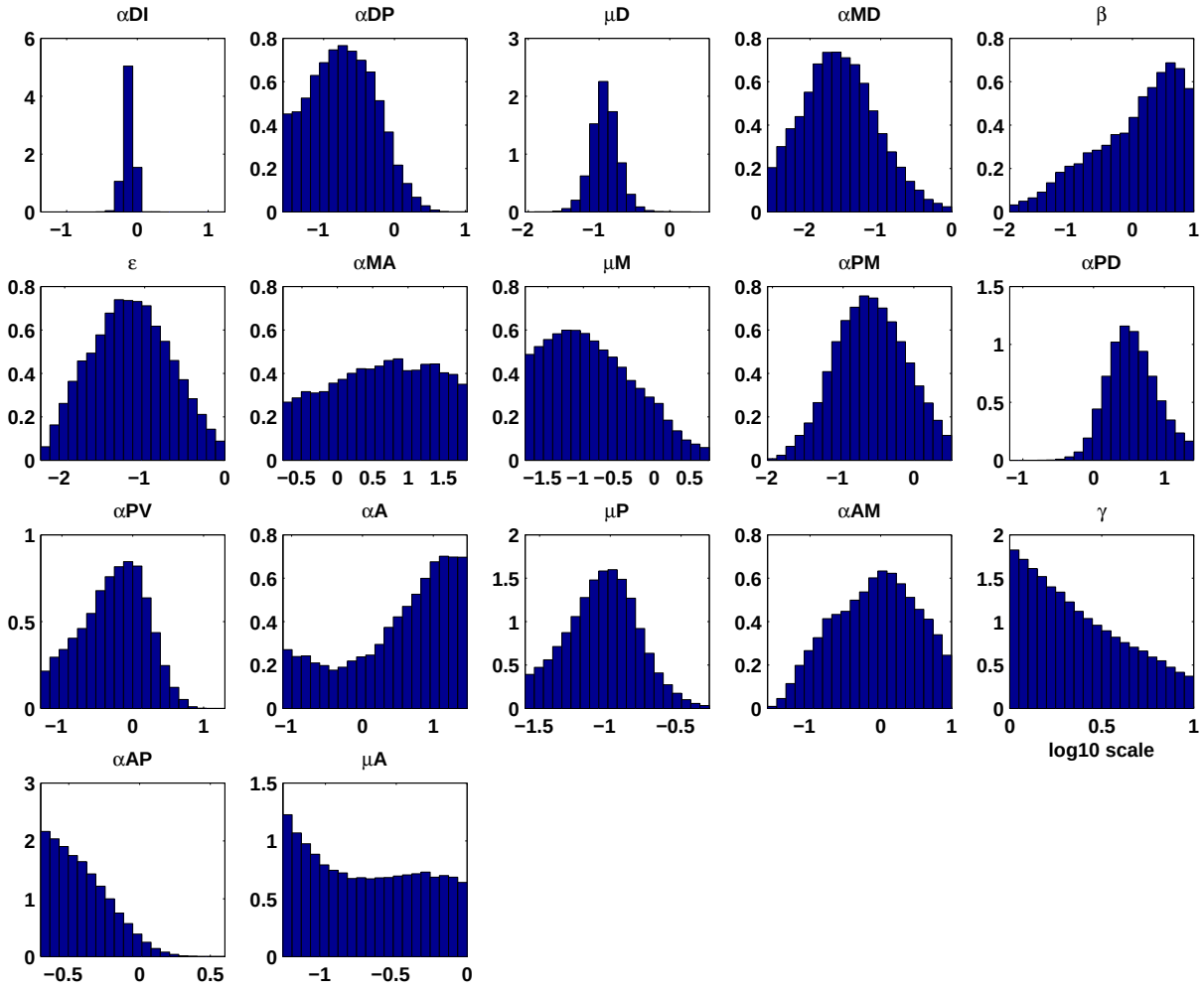

Figure 36: Marginal parameter distributions for the ensemble model. The X-axis is in log10 scale.

The trajectory statistics for the ensemble model is shown in Figure 37 as not as individual trajectories, but as percentile levels of the distribution of the variable versus time. For example, if the 75th percentile curve of A has value A_0 at time t_0, then 75% of all trajectories in the ensemble have $A(t_0) \leqslant A_0$. As see in Figure 37, the variances in the values of both pro- and anti-inflammatories are approximately constant in time. The large increase in the variance of M is a result of having no constraint on the maximum value of M. Again, the distributions reflect the inability of the model to capture the decrease in IL-10 for the voice

rest treatment.

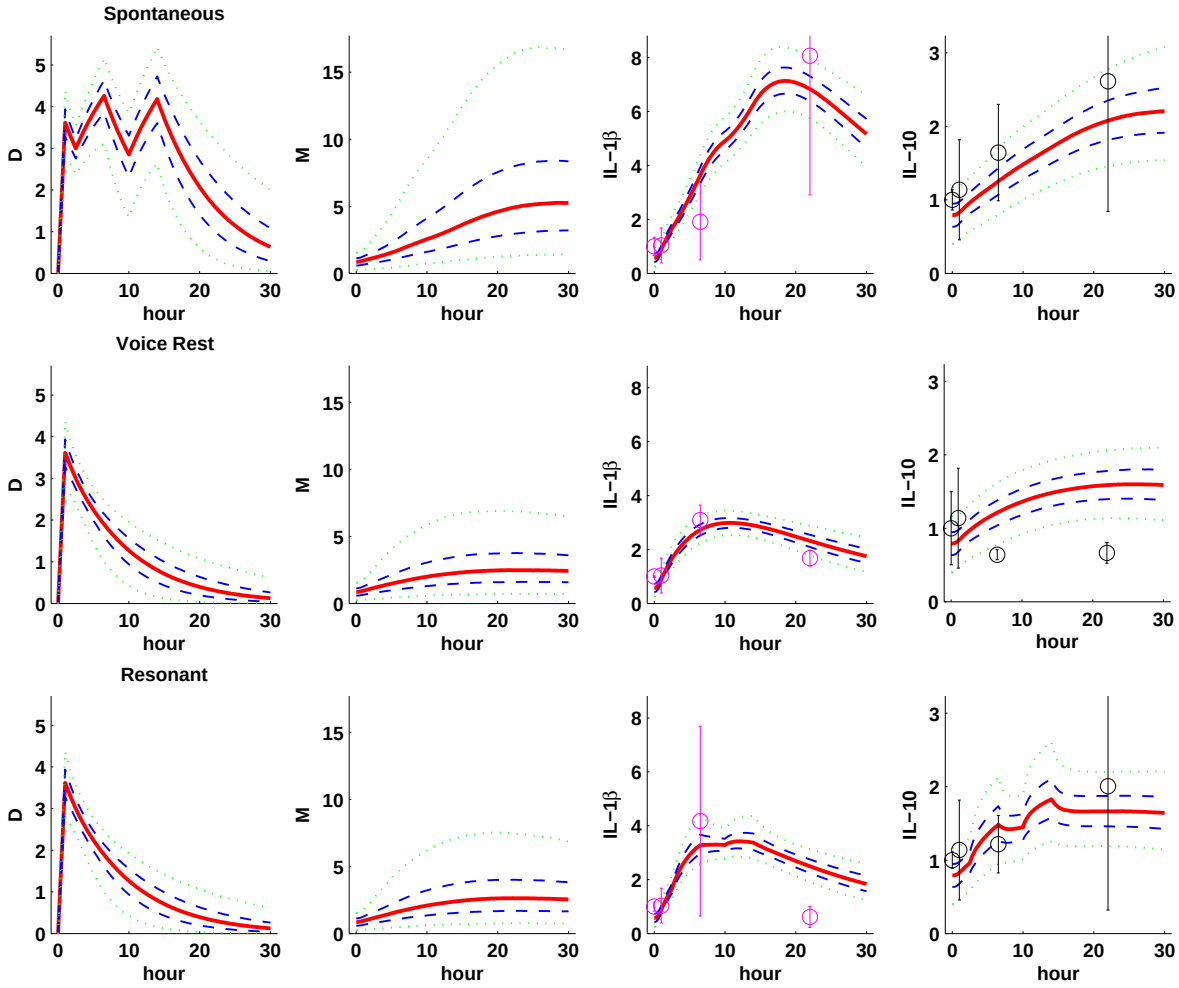

Figure 37: Trajectory statistics of the ensemble model. For each variable, the percentile of the resulting distribution is shown at each time point: 95th percentile (top green), 75th percentile (top blue), 50th percentile (red), 25th percentile (bottom blue), and 5th percentile (bottom green).

3.3.5.2 Sensitivity Analysis and Parameter Reduction By sorting the diagonal elements of the covariance matrix C of $\bar{G}$, we have the most to the least sensitive parameter in the following order: 1, 9, 3, 17, 14, 10, 13, 4, 15, 12, 5, 8, 7, 16, 2, 11, 6. We assume that two parameters are highly correlated if their absolute value of the covariance coefficient is great than 0.95. Parameter 1 is the most sensitive one - it is highly correlated with 9 and

10; go to the next most sensitive parameter - 3 is highly correlated with 2 and 13; go to next sensitive parameter - 17 is highly correlated with 14, 15, and 16; then parameter 4 is highly correlated with parameter 5 and 12; then parameter 8 is highly correlated with 8; and 11 is not highly correlated with any of the others. So the reduced parameter set are: 1, 3, 4, 8, 11 and 17, corresponding to α_{DI}, μ_D, α_{MD}, μ_M, α_{PV} and μ_A. We have also tried the singular value decomposition to calculate the sensitivities, but it doesn't perform as good as this one.

The sensitivity-dependent correlation matrix correlation containing the correlation coefficients of the 17 parameters is in Figure 38. Red represents positive highly correlated, blue represents negative highly correlated, and the other colors are the intermediate ones.

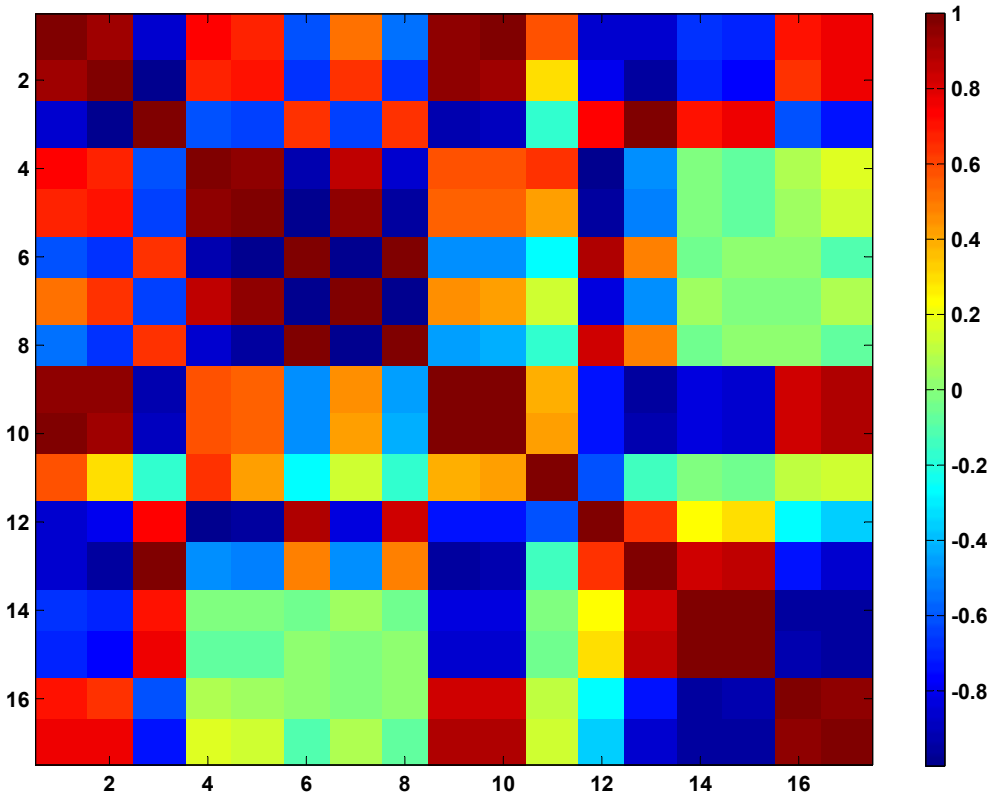

Figure 38: Correlation coefficient matrix. The red color means highly positive correlated and the blue color means highly negative correlated.

3.3.5.3 Pre-inflamed Group We mentioned before that there are two patients who received the treatment in voice rest having high initial value of IL-1β. These two patients are defined as pre-inflamed ones. The baseline values of IL-1β and IL-10 are normalized by the average of other two patients' baselines in voice rest group. The initial value of this group is $(D, M, P, A) = (0, 1, 11.33, 0.76)$. The full parameter sets got from the sampling method above are applied to the pre-inflamed group. We also run the statistical trajectories of this group.

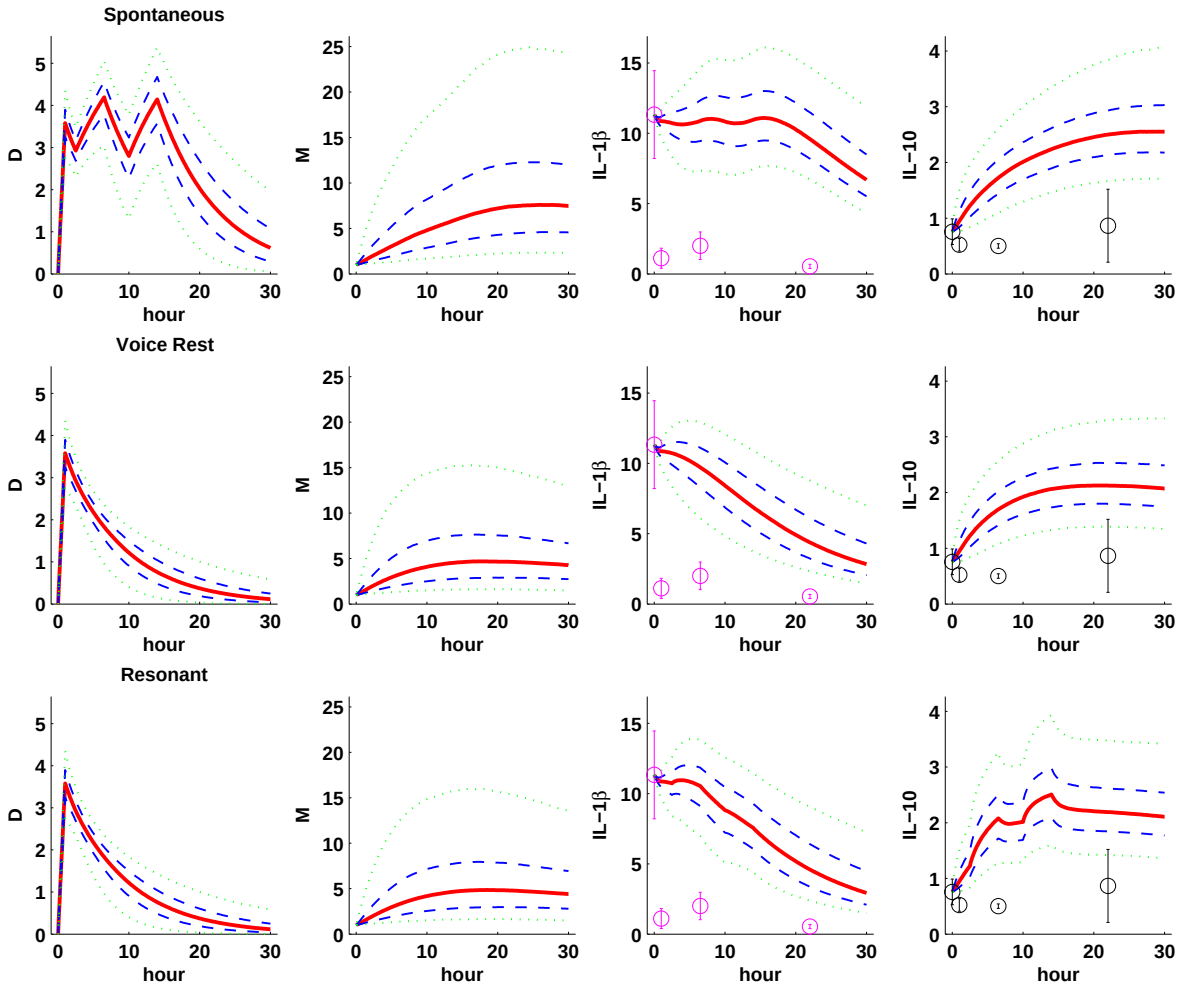

Figure 39: Trajectory statistics of the ensemble model of the pre-inflamed group which only received treatment in voice rest. For each variable the percentile of the resulting distribution is shown at each time point: 95th percentile (top green), 75th percentile (top blue), 50th percentile (red), 25th percentile (bottom blue), and 5th percentile (bottom green).

The two patients in the pre-inflamed group received the treatment of voice rest. Data points of IL-1β and IL-10 in Figure 39 are all measured under the treatment of voice rest. The damage behaviors of the pre-inflamed patients are almost the same as those of the normal patients. Thus we could see that pre-inflammation would not cause extra damage during the phonation loading. The magnitudes of macrophage, IL-1β and IL-10 in pre-inflamed patients are much higher than the normal patients in corresponding treatment group. IL-1β has small peak in spontaneous and decreases almost consistently in the other two treatments. The IL-10 increases in all three treatments and peaks at similar levels. From the statistical trajectory, we can see that the best therapy to heal from vocal fold inflammation for those pre-inflamed people is also voice rest or treatment in resonant voice.

3.3.5.4 Individual Fit and prediction There is one special subject who receives the 3 different treatments at different time. We fit this one through the reduced parameter set. We only input the data of first three time points and try to predict the last one. Figure 40 shows one of the optimal fits.

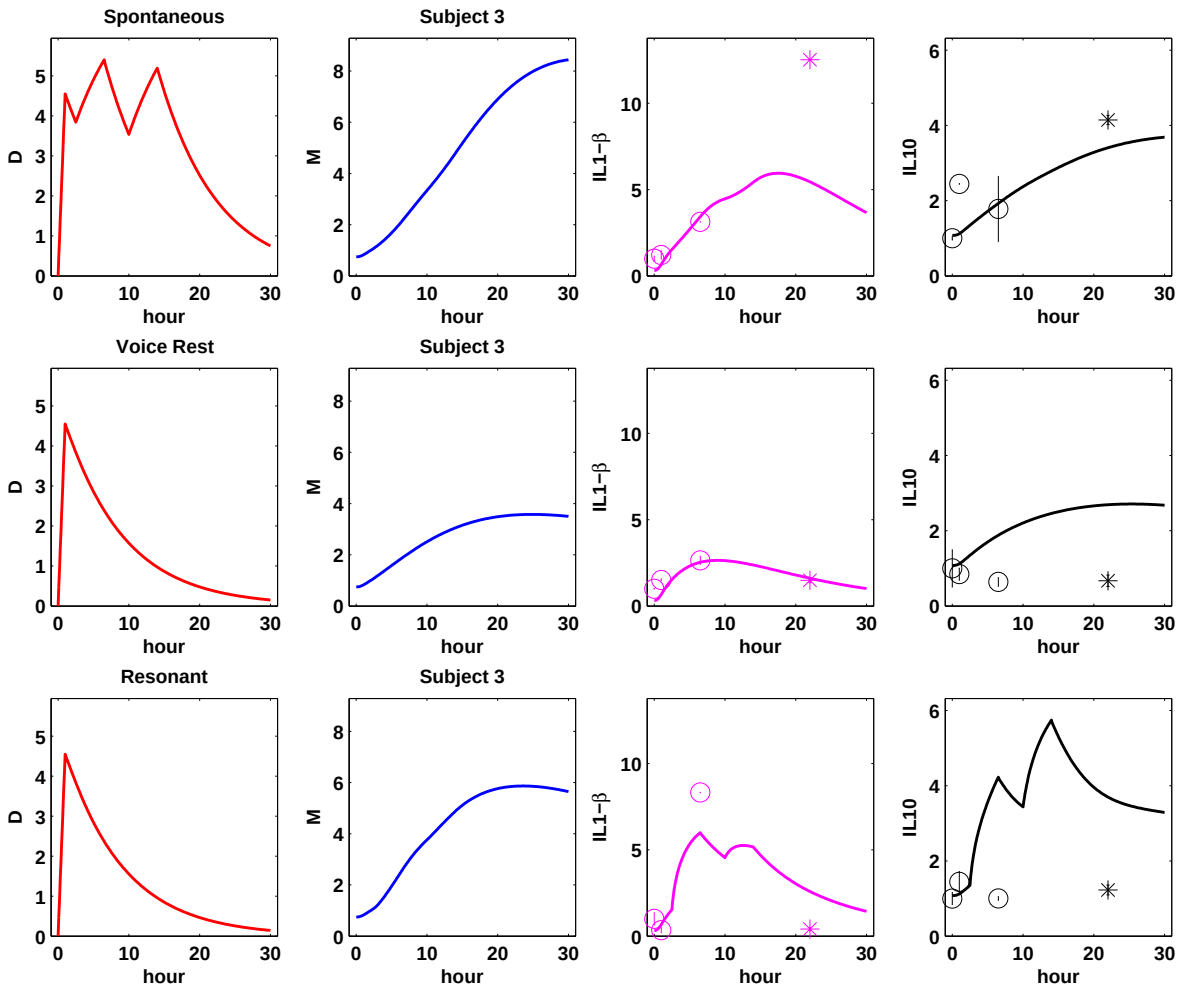

Figure 40: Optimal parameter fit and prediction for individual. The parameter values are showed in the last column of Table 5.

In spontaneous, the model tries to predict the peak of IL-1β. The predicted peak time is right but the magnitude is not as high as the real data. Though the second point of IL-10 is not assimilated in the model, but the peak of IL-10 is predicted. In voice rest, the model assimilates the first three data points of IL-1β very well and the prediction of the fourth agrees with the real data point. For the IL-10, again, the model hasn't captured the decrease of IL-10. In the resonant voice, the model's assimilation and prediction of IL-1β is not perfect, but acceptable. Then again, the decrease of IL-10 is not captured.

Parameter values in Figure 40 are showed in the last column of Table 5. Parameters in red are the new optimal values of the reduced set and black are the fixed less sensitive parameters

obtained from the average in parallel samplings. Table 5 also lists the average and standard derivation of parameters in parallel sampling. We can see except one parameter α_{PV}, all the other new optimal parameter values are within the standard derivation of the average values in the full model. α_{PV} is an exception might because this parameter characterizes the effect vibration stress in pro- and anti-inflammatories. Each individual may react quite differently to a same treatment therapy.

No	Parameter	Average	SD	Optimal
1	α_{DI}	0.7631	0.1172	0.9474
2	α_{DP}	0.3194	0.3974	0.3194
3	μ_D	0.133	0.0681	0.1213
4	α_{MD}	0.0566	0.0943	0.0288
5	β	2.5262	2.5955	2.5262
6	ε	0.1239	0.1537	0.1239
7	α_{MA}	11.7458	15.6568	11.7458
8	μ_M	0.3912	0.7383	0.3479
9	α_{PM}	0.4534	0.5541	0.4534
10	α_{PD}	4.8495	4.3919	4.8495
11	α_{PV}	0.8571	0.8845	8.1921
12	α_A	8.0904	8.2102	8.0904
13	μ_P	0.1037	0.0628	0.1037
14	α_{AM}	1.7211	2.0583	1.7211
15	γ	2.9009	2.0393	2.9009
16	α_{AP}	0.4386	0.2520	0.4386
17	μ_A	0.2909	0.2524	0.0953

Table 5: Parameters in red are the new optimal values of the reduced set and black are the fixed parameters obtained from the averages of the whole parameter set in parallel samplings. The values in the other two columns are average and standard derivation of each parameter in parallel sampling.

3.4 CONCLUSION

We have employed mathematical modeling to explain differences in trajectories of patients with vocal fold inflammation subject to individualized treatments. We have designed an ensemble model representing the parameter variability of the patient population and found that it mostly agrees with the data. We conducted the sensitivity analysis in order to find the global and local parameters and then fix the global parameters and optimize the local ones for individuals. The model was not very successful in fitting to individual patients trajectories.

The first limitation of this model is that it is simplified to only 4 variables. More complex models can be constructed, but ultimately, the limited data set and the high variance restricted our ability to fit such models. The inclusion of the fibroblast and collagen into the model will result in a better understanding of the healing process. The second limitation of the current model is its simplified account of the loading conditions, represented by vibratory or impact stress variables. We need to have a more qualitative understanding of the magnitude and effect of the two stresses. The third limitation is the fact that since our reduction methods resulted in missing IL-10 trajectories, rather than isolated data points, it might be more appropriate to adopt global methods, rather than local methods in this situation. Such methods typically results is less parsimonious reduction in the number of parameters and thus greater ability to fit data.

Our future goal is to develop a method for proposing a patient specific vocal exercise schedule or rest program that optimizes tissue healing in cases of both acute and chronic phonotrauma.

BIBLIOGRAPHY

[1] M. Adib-Conquy and J.M. Cavaillon. Compensatory anti-inflammatory response syndrome. *Thromb.Haemost.*, 101(1):36–47, January 2009.

[2] P.B. Bitterman, M.D. Wewers, S.I. Rennard, S. Adelberg, and R.G. Crystal. Modulation of alveolar macrophage-driven fibroblast proliferation by alternative macrophage mediators. *J.Clin.Invest*, 77(3):700–708, March 1986.

[3] G.A. Bocharov and A.A. Romanyukha. Mathematical model of antiviral immune response. iii. influenza a virus infection. *J Theor.Biol.*, 167(4):323–360, April 1994.

[4] C. Bogdan, J. Paik, Y. Vodovotz, and C. Nathan. Contrasting mechanisms for suppression of macrophage cytokine release by transforming growth factor-beta and interleukin-10. *J.Biol.Chem.*, 267(32):23301–23308, November 1992.

[5] C. Bogdan, Y. Vodovotz, and C. Nathan. Macrophage deactivation by interleukin 10. *J.Exp.Med.*, 174(6):1549–1555, December 1991.

[6] D.R. Boone and S.C. McFarlane. *The Voice and Voice Therapy (5th ed.)*. 1994.

[7] R.C. Branski, S.S. Barbieri, B.B. Weksler, B. Saltman, P. Krishna, D.H. Kraus, N.V. Broadbelt, J. Chen, D.P. Poppas, and D. Felsen. Effects of transforming growth factor-beta1 on human vocal fold fibroblasts. *Ann.Otol.Rhinol.Laryngol.*, 118(3):218–226, March 2009.

[8] Y. Cao, D.T. Gillespie, and L.R. Petzold. Efficient step size selection for the tau-leaping simulation method. *J Chem Phys*, 124(4):044109–, January 2006.

[9] Y. Cao, D.T. Gillespie, and L.R. Petzold. Adaptive explicit-implicit tau-leaping method with automatic tau selection. *J Chem Phys*, 126(22):224101–, June 2007.

[10] Y. Cao, H. Li, and L. Petzold. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *J Chem Phys*, 121(9):4059–4067, September 2004.

[11] H. Chanteux, A.C. Guisset, C. Pilette, and Y. Sibille. Lps induces il-10 production by human alveolar macrophages via mapkinases- and sp1-dependent mechanisms. *Respir.Res.*, 8:71–, 2007.

[12] C.C. Chow, G. Clermont, R. Kumar, C. Lagoa, Z. Tawadrous, D. Gallo, B. Betten, J. Bartels, G. Constantine, M.P. Fink, T.R. Billiar, and Y. Vodovotz. The acute inflammatory response in diverse shock states. *Shock*, 24(1):74–84, July 2005.

[13] G. Clermont, J. Bartels, R. Kumar, G. Constantine, Y. Vodovotz, and C. Chow. In silico design of clinical trials: a method coming of age. *Crit Care Med.*, 32(10):2061–2070, October 2004.

[14] R. Colton and J.K. Casper. *Understanding Voice Problems: A Physiological Perspective for Diagnosis and Treatment (2nd ed.)*. 1996.

[15] E. Cuthill and J. McKee. Reducing the bandwidth of sparse symmetric matrices. In *Proc. 24th Nat. Conf. ACM*, pages 157–172, 1969.

[16] S. Daun, J. Rubin, Y. Vodovotz, A. Roy, R. Parker, and G. Clermont. An ensemble of models of the acute inflammatory response to bacterial lipopolysaccharide in rats: results from parameter space reduction. *J Theor.Biol.*, 253(4):843–853, August 2008.

[17] J. Day, A. Friedman, and L.S. Schlesinger. Modeling the immune rheostat of macrophages in the lung in response to infection. *Proc.Natl.Acad.Sci.U.S.A*, 106(27):11246–11251, July 2009.

[18] J. Day, J. Rubin, Y. Vodovotz, C.C. Chow, A. Reynolds, and G. Clermont. A reduced mathematical model of the acute inflammatory response ii. capturing scenarios of repeated endotoxin administration. *J Theor.Biol.*, 242(1):237–256, September 2006.

[19] C.A. Dinarello. Induction of interleukin-1 and interleukin-1 receptor antagonist. *Semin.Oncol.*, 24(3 Suppl 9):S9–S9, June 1997.

[20] C.A. Dinarello. Proinflammatory cytokines. *Chest*, 118(2):503–508, August 2000.

[21] C.R. Doering, K.V. Sargsyan, and L.M. Sander. Extinction times for birth-death processes: exact results, continuum asymptotics, and the failure of the fokker-planck approximation. *MULTISCALE MODEL.SIMUL.*, 3(2):283–299, 2005.

[22] M.R. Duncan and B. Berman. Stimulation of collagen and glycosaminoglycan production in cultured human adult dermal fibroblasts by recombinant human interleukin 6. *J.Invest Dermatol.*, 97(4):686–692, October 1991.

[23] M.R. Duncan, K.S. Frazier, S. Abramson, S. Williams, H. Klapper, X. Huang, and G.R. Grotendorst. Connective tissue growth factor mediates transforming growth factor beta-induced collagen synthesis: down-regulation by camp. *FASEB J.*, 13(13):1774–1786, October 1999.

[24] M.I. Dykman, I.B. Schwartz, and A.S. Landsman. Disease extinction in the presence of non-gaussian noise. 2008.

[25] D.J. Earl and M.W. Deem. Parallel tempering: theory, applications, and new perspectives. *Phys Chem Chem Phys*, 7(23):3910–3916, December 2005.

[26] Johan Elf, Per Lotstedt, and Paul Sjoberg. Problems of high dimension in molecular biology. In *Proceedings of the 19th GAMM-Seminar, Leipzig*, pages 21–30, 2003.

[27] W. Feller. *An Introduction to Probability Theory and Its Applications*. 1968.

[28] A Gelman and D.B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7:457–472, 1992.

[29] M.A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J.Phys.Chem.A*, 104(9):1876–1889, March 2000.

[30] D.T. Gillespie. A general method of numerically simulating the stochastic time evolution of coupled chemical reactions. *J.Computational Phys.*, 22:403–434, 1976.

[31] D.T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J.Phys.Chem.*, 81(25):2340–2361, 1977.

[32] D.T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A*, 188(1-3):404–425, September 1992.

[33] D.T. Gillespie. The chemical langevin equation. *J Chem Phys*, 113(1):–, 2000.

[34] D.T. Gillespie. Stochastic simulation of chemical kinetics. *Annu.Rev.Phys Chem*, 58:35–55, 2007.

[35] D.T. Gillespie and L. Petzold. Improved leap-size selcetion for accelerated stochastic simulation. *J Chem Phys*, 119(16):–, 2003.

[36] S. Gray and I. Titze. Histologic investigation of hyperphonated canine vocal cords. *Ann.Otol.Rhinol.Laryngol.*, 97(4 Pt 1):381–388, July 1988.

[37] H.E. Gunter. A mechanical model of vocal-fold collision with high spatial and temporal resolution. *J Acoust.Soc.Am.*, 113(2):994–1000, February 2003.

[38] H.E. Gunter. Modeling mechanical stresses as a factor in the etiology of benign vocal fold lesions. *J Biomech.*, 37(7):1119–1124, July 2004.

[39] B. Hancioglu. *Mathematical Modeling of Virus Dynamics in Immunology*. PhD thesis, University of Pittsbugh, 2007.

[40] B. Hancioglu, D. Swigon, and G. Clermont. A dynamical model of human immune response to influenza a virus infection. *J.Theor.Biol.*, 246(1):70–86, May 2007.

[41] W.K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[42] Shin ichi Tamura and Takeshi Kurata. Defense mechanisms against influenza virus infection in the respiratory tract mucosa. *Jpn J Infect Dis*, 57(6):236–47, 2004.

[43] J.J. Jiang, C.E. Diaz, and D.G. Hanson. Finite element modeling of vocal fold vibration in normal phonation and hyperfunctional dysphonia: implications for the pathogenesis of vocal nodules. *Ann.Otol.Rhinol.Laryngol.*, 107(7):603–610, July 1998.

[44] M.M. Johns. Update on the etiology, diagnosis, and treatment of vocal fold nodules, polyps, and cysts. *Curr.Opin.Otolaryngol.Head Neck Surg.*, 11(6):456–461, December 2003.

[45] M. Jordana, B. Sarnstrand, P.J. Sime, and I. Ramis. Immune-inflammatory functions of fibroblasts. *Eur.Respir.J.*, 7(12):2212–2222, December 1994.

[46] A. Kamenev and B. Meerson. Extinction of an infectious disease: A large fluctuation in a nonequilibrium system. *Physical Review e*, 77(6):–, June 2008.

[47] D. Kraig. Introduction to stochastic simulation with the gillespie method, 2005.

[48] R. Kumar, G. Clermont, Y. Vodovotz, and C.C. Chow. The dynamics of acute inflammation. *J Theor.Biol.*, 230(2):145–155, September 2004.

[49] J.K. Kutty and K. Webb. Vibration stimulates vocal mucosa-like matrix expression by hydrogel-encapsulated fibroblasts. *J Tissue Eng Regen.Med.*, 4(1):62–72, January 2010.

[50] P.D. Leenheer and H.L. Smith. Virus dynamics: a global analysis. *SIAM J.APPL.MATH.*, 63(4):1313–1327, 2003.

[51] N.Y. Li. *Biosimulation of vocal fold inflammation and healing.* PhD thesis, University of Pittsburgh, 2009.

[52] N.Y. Li, K. Verdolini, G. Clermont, Q Mi, P.A. Hebda, and Y. Vodovotz. Simulating acute phonotrauma: a comparison of agent-based and equation-based models. In *the 5th International Conference on Voice Physiology and Biomechanics*, pages –, 2006.

[53] N.Y. Li, K. Verdolini, G. Clermont, Q. Mi, E.N. Rubinstein, P.A. Hebda, and Y. Vodovotz. A patient-specific in silico model of inflammation and healing tested in acute vocal fold injury. *PLoS.One.*, 3(7):e2789–, 2008.

[54] X. Lim, D.M. Bless, A. Munoz-Del-Rio, and N.V. Welham. Changes in cytokine signaling and extracellular matrix production induced by inflammatory factors in cultured vocal fold fibroblasts. *Ann.Otol.Rhinol.Laryngol.*, 117(3):227–238, March 2008.

[55] X. Lim, I. Tateya, T. Tateya, A. Munoz-Del-Rio, and D.M. Bless. Immediate inflammatory response and scar formation in wounded vocal folds. *Ann.Otol.Rhinol.Laryngol.*, 115(12):921–929, December 2006.

[56] J.E. Losa Garcia, F.M. Rodriguez, M.R. Martin de Cabo, M.J. Garcia Salgado, J.P. Losada, L.G. Villaron, A.J. Lopez, and J.L. Arellano. Evaluation of inflammatory cytokine secretion by human alveolar macrophages. *Mediators.Inflamm.*, 8(1):43–51, 1999.

[57] Y. Luo, J.B. Kobler, S.M. Zeitels, and R. Langer. Effects of growth factors on extracellular matrix production by vocal fold fibroblasts in 3-dimensional culture. *Tissue Eng*, 12(12):3365–3374, December 2006.

[58] E.P. Ma and E.M. Yiu. Voice activity and participation profile: assessing the impact of voice disorders on daily activities. *J Speech Lang Hear.Res.*, 44(3):511–524, June 2001.

[59] J. Ma, T. Chen, J. Mandelin, A. Ceponis, N.E. Miller, M. Hukkanen, G.F. Ma, and Y.T. Konttinen. Regulation of macrophage activation. *Cell Mol.Life Sci.*, 60(11):2334–2346, November 2003.

[60] Donald A. McQuarrie. Stochastic approach to chemical kinetics. *J.Appl.Probability*, 4:413–478, 1967.

[61] S. Merrill. *Modeling the interaction of HIV with the cells of the immune system, in Mathematical and Statistical Approaches to AIDS Epidemiology*. Springer-Verlag, New York, 1989.

[62] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092, 1953.

[63] K.W. Moore, Malefyt R. de Waal, R.L. Coffman, and A. O'Garra. Interleukin-10 and the interleukin-10 receptor. *Annu.Rev.Immunol.*, 19:683–765, 2001.

[64] I. Nasell. The quasi-stationary distribution of the closed endemic sis model. *Adv.in Appl.Probab.*, 28(3):895–932, 1996.

[65] I. Nasell. On the time to extinction in recurrent epidemics. *J.R.Stat.Soc.Ser.B Stat.Methodol.*, 61(2):309–330, 1999.

[66] I. Nasell. Extinction and quasi-stationarity in the verhulst logistic model. *J.Theor.Biol.*, 211(1):11–27, July 2001.

[67] T.J. Newman, J.B. Ferdy, and C. Quince. Extinction times and moment closure in the stochastic logistic process. *Theor.Popul.Biol.*, 65(2):115–126, March 2004.

[68] M.A. Nowak and C.R.M. Bangham. Population dynamics of immune responses to persistent viruses. *Science*, 272(5258):74–79, April 1996.

[69] M.A. Nowak and May R.M. *Virus Dynamics*. Oxford University Press, New York, 2000.

[70] T. Ohno, S. Hirano, and B. Rousseau. Gene expression of transforming growth factor-beta1 and hepatocyte growth factor during wound healing of injured rat vocal fold. *Laryngoscope*, 119(4):806–810, April 2009.

[71] A.S. Perelson. Modelling viral and immune system dynamics. *Nat.Rev.Immunol.*, 2(1):28–36, January 2002.

[72] A.S. Perelson and D.E. Kirschner. Dynamics of hiv infection of cd4+ t cells. *Math.Biosci.*, 114:81–125, 1993.

[73] K.L. Peterson, K. Verdolini-Marston, J.M. Barkmeier, and H.T. Hoffman. Comparison of aerodynamic and electroglottographic parameters in evaluating clinically relevant voicing patterns. *Ann.Otol.Rhinol.Laryngol.*, 103(5 Pt 1):335–346, May 1994.

[74] L.O. Ramig and K. Verdolini. Treatment efficacy: voice disorders. *J Speech Lang Hear.Res.*, 41(1):S101–S116, February 1998.

[75] M. Rathinam, L. Petzold, Y. Cao, and D.T. Gillespie. Stiffness in stochastic chemically reacting systems: the implicit tau-leaping method. *J Chem Phys*, 119(24):–, 2010.

[76] A. Reynolds, J. Rubin, G. Clermont, J. Day, Y. Vodovotz, and Ermentrout G. Bard. A reduced mathematical model of the acute inflammatory response: I. derivation of model and analysis of anti-inflammation. *J Theor.Biol.*, 242(1):220–236, September 2006.

[77] P Sjoberg, P Lotstedt, and J Elf. Fokker-planck approximation of the master equation in molecular biology. *Computing and Visualization in Science*, 12(1):37–50, 2007.

[78] William J. Stewart. *Introduction to the numerical solution of markov chains*. 1994.

[79] M.J. Thomassen, L.T. Divis, and C.J. Fisher. Regulation of human alveolar macrophage inflammatory cytokine production by interleukin-10. *Clin.Immunol.Immunopathol.*, 80(3 Pt 1):321–324, September 1996.

[80] I.R. Titze. Mechanical stress in phonation. *J Voice*, 8(2):99–105, June 1994.

[81] I.R. Titze, J.G. Svec, and P.S. Popolo. Vocal dose measures: quantifying accumulated vibration exposure in vocal fold tissues. *J Speech Lang Hear.Res.*, 46(4):919–932, August 2003.

[82] O.A. van Herwaarden. Stochastic epidemics: the probability of extinction of an infectious disease at the end of a major outbreak. *J Math.Biol.*, 35(7):793–813, August 1997.

[83] M. Vellela and H. Qian. A quasistationary analysis of a stochastic chemical reaction: Keizer's paradox. *Bull.Math.Biol.*, 69(5):1727–1746, July 2007.

[84] K. Verdolini, D.G. Druker, P.M. Palmer, and H. Samawi. Laryngeal adduction in resonant voice. *J Voice*, 12(3):315–327, September 1998.

[85] K. Verdolini, M.M. Hess, I.R. Titze, W. Bierhals, and M. Gross. Investigation of vocal fold impact stress in human subjects. *J Voice*, 13(2):184–202, June 1999.

[86] K. Verdolini and L.O. Ramig. Review: occupational risks for voice problems. *Logoped.Phoniatr.Vocol.*, 26(1):37–46, 2001.

[87] P.F. Verhulst. Notice sur la loi que la population suit dans son accroisement. *Corr.Math.et Phys.*, X:113–121, 1838.

[88] Y. Vodovotz. Deciphering the complexity of acute inflammation using mathematical models. *Immunol.Res.*, 36(1-3):237–245, 2006.